

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:26:41 ; Search time 18 Seconds
(without alignments)
717.411 Million cell updates/sec

Title: US-10-015-385A-194

Perfect score: 1374

Sequence: 1 MGSLFELLCLVGLSQATP.....GVYTYICKYVDWIRMRNN 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1374	100.0	248	1	KLK8_HUMAN
2	630.5	45.9	260	1	KLK8_HUMAN
3	622.5	45.3	260	1	NRPN_MOUSE
4	621.5	45.2	260	1	NRPN_RAT
5	618.5	45.0	250	1	KLK8_HUMAN
6	599	43.6	256	1	KLK8_HUMAN
7	569.5	41.4	248	1	TRY1_CHICK
8	569.5	41.4	251	1	KLK8_HUMAN
9	568.5	41.4	250	1	KLK9_HUMAN
10	567	41.3	277	1	KLK9_HUMAN
11	566	41.2	248	1	TRY2_CHICK
12	565	41.1	248	1	TRY1_CHICK
13	564.5	41.1	293	1	KLK5_HUMAN
14	564	41.0	276	1	KLK4_HUMAN
15	563	41.0	244	1	TRY2_XENLA
16	561	40.8	231	1	TRY1_PIG
17	561	40.8	243	1	TRY1_BOVIN
18	561	40.8	244	1	KLK6_HUMAN
19	555	40.4	243	1	TRY1_XENLA
20	555	40.4	253	1	KLK7_HUMAN
21	553	40.2	246	1	TRY2_MOUSE
22	550	40.0	247	1	TRY2_BOVIN
23	546	39.7	247	1	TRY1_CANFA
24	545	39.7	246	1	TRY1_CANFA
25	544.5	39.6	247	1	TRY3_RAT
26	543	39.5	246	1	TRY1_RAT
27	540	39.3	238	1	TRY3_SALSA
28	533.5	38.8	247	1	TRY4_RAT
29	533	38.8	246	1	TRY2_RAT
30	527.5	38.4	261	1	KLK7_RAT
31	521.5	38.0	246	1	KLK_PIG
32	521.5	38.0	261	1	KLK6_MOUSE
33	519	37.8	246	1	TRYA_RAT

34	518	37.7	246	1	TRYB_RAT
35	518	37.7	261	1	KLK8_RAT
36	513.5	37.4	231	1	TRY2_SALSA
37	513.5	37.4	239	1	KLK2_CAVPO
38	510.5	37.2	261	1	KLK3_MOUSE
39	508	37.0	242	1	TRY1_SALSA
40	505	36.8	260	1	ESTA_CANFA
41	503	36.6	263	1	KLK8_PRANA
42	502	36.5	259	1	KLK8_RAT
43	501	36.5	247	1	TRY2_HUMAN
44	499.5	36.4	241	1	TRYX_GADMO
45	498	36.2	254	1	KLK4_HUMAN

ALIGNMENTS

RESULT 1	KLK8_HUMAN	STANDARD;	PRT;	248 AA.
AC	Q9UKR0; Q9UKR1;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Kallikrein 12 precursor (EC 3.4.21.-) (Kallikrein-like protein 5) (KLK-L5).			
GN	KLK12 OR KLK15.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	MEDLINE=20118156; PubMed=10652563;			
RX	Yousef G.M., Luo L.-Y., Diamandis E.P.;			
RA	"Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4";			
RT	Anticancer Res. 19:2843-2852(1999).			
RL	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RA	Yousef G.M., Magkara A., Scorilas A., Diamandis E.P.;			
RT	"Cloning of new alternatively spliced forms of the kallikrein-like gene 5 (KLK-L5).";			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.			
[3]	SEQUENCE FROM N.A. (ISOFORM 1).			
RP	MEDLINE=20510030; PubMed=11054574;			
RX	Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,			
RA	Moss P., Paepker B., Wang K.;			
RT	"Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";			
RL	Gene 257:119-130(2000).			
[4]	SEQUENCE FROM N.A. (ISOFORM 2).			
RP	Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,			
RA	Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,			
RT	Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Barnes J.,			
RL	Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,			
RA	Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefficient J.,			
RA	Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,			
RA	Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,			
RA	Olsen A.S., Carraro A.V.;			
RT	"Sequence analysis of chromosome 19q13.4";			
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: Secreted (Probable).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=1;			
CC	IsoId=Q9UKR0-1; Sequence=Displayed;			
CC	Name=2;			
CC	IsoId=Q9UKR0-2; Sequence=VSP_005403;			
CC	-1- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/> or send an email to license@ebi.ac.uk).

```

DR EMBL; AF135025; AAD26426.2; -
DR EMBL; AF135025; AAF06065.1; -
DR EMBL; AF243527; AAG33365.1; -
DR EMBL; AC011473; AAG32258.1; -
DR HSSP; PO0763; 1DP0.
DR MEROPS; S01.020; -.
DR Genew; HGNC:6360; KLK12.
DR MIM; 605539; -. C:extracellular; NAS.
DR GO; GO:0005576; C:serine-type endopeptidase activity; NAS.
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase, Serine protease, Glycoprotein, Signal,
KW Alternative splicing.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 248 KALLIKREIN 12.
FT ACT_SITE 62 62 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 28 161 BY SIMILARITY.
FT DISULFID 47 63 BY SIMILARITY.
FT DISULFID 133 235 BY SIMILARITY.
FT DISULFID 140 206 BY SIMILARITY.
FT DISULFID 172 186 BY SIMILARITY.
FT DISULFID 196 222 BY SIMILARITY.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 236 248 KVDWIRIMIRNN -> NSTVLVGLGTSWNFNSQPF (in isoform 2).
FT /FTID:VSP_005403.
SQ SEQUENCE 248 AA; 26733 MW; BB473E98F8BAF703 CRC64;
Query Match 100.0%; Score 1374; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.7e-105;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLCVLGLSQAATPKIPNGTECGRNSQPQWQVGLFEGTSRCGGVLIDHRWLTA 60
Db 1 MGLSIFLLCVLGLSQAATPKIPNGTECGRNSQPQWQVGLFEGTSRCGGVLIDHRWLTA 60
QY 61 AHCSGRYVRLGEHSLSDWTEQIRHSGFSTHYPGILGASTSHEDLRLRLPVRV 120
Db 61 AHCSGRYVRLGEHSLSDWTEQIRHSGFSTHYPGILGASTSHEDLRLRLPVRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGTINHPNPFDPDLQCLNLSIVSHATCHGVYPGRI 180
Db 121 TSSVQPLPLPNDCATAGTECHVSGWGTINHPNPFDPDLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGGVPQDACCQGGSLVCGVQLGLVSWGSGVPCQGDGIPGYTIVICKYVDW 240
Db 181 TSNMVCAGGVPQDACCQGGSLVCGVQLGLVSWGSGVPCQGDGIPGYTIVICKYVDW 240
QY 241 IRIMIRNN 248
Db 241 IRIMIRNN 248

```

```

RESULT 2
ID KLB8 HUMAN STANDARD; PRT; 260 AA.
AC OG0259; Q9HCB3; Q9UIL9; Q9UQ47;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neuropsin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Ovasin) (Serine
DE protease TADG-14) (Tumor-associated differentially expressed gene-14
DE protein).
GN KLB8 OR PRS19 OR TADG14 OR NRPN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hippocampus;
RX MEDLINE=98372070; PubMed=9714609;
RA Yoshida S., Taniguchi M., Hirata A., Shiosaka S.;
RT "Sequence analysis and expression of human neuropsin cDNA and gene.";
RL Gene 213:9-16(1998).
[2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=98203457; PubMed=10102990;
RA Mitsui S., Tsuruoka N., Yamashiro K., Nakazato H., Yamaguchi N.;
RT "A novel form of human neuropsin, a brain-related serine protease, is
RT generated by alternative splicing and is expressed preferentially in
RT human adult brain.";
RL Eur. J. Biochem. 260:627-634(1999).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Ovary;
RX MEDLINE=99413504; PubMed=10485494;
RA Underwood L.J., Tanimoto H., Wang Y., Shigemasa K., Farmlay T.H.,
RA O'Brien T.J.;
RT "Cloning of tumor-associated differentially expressed gene-14, a novel
RT serine protease overexpressed by ovarian carcinoma.";
RL Cancer Res. 59:4435-4439(1999).
[4]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RA Gan L., Gellinas R., Gown A.M., Moss P., Smith R., Wang K.;
RT "Molecular cloning and characterization of a novel serine protease,
RT ovasin, a potential molecular marker for ovarian carcinomas.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA Moss P., Paepker B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
[6]
RN SEQUENCE OF 1-164 FROM N.A. (ISOFORM 1).
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,
RA Dangnan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Suggested to be involved in kindling epileptogenesis and
CC hippocampal plasticity.
CC -! CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.
CC -! SUBCELLULAR LOCATION: Secreted.
CC -! ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```



```

CC CC      Isoid=060259-1; Sequence=Displayed;
CC CC      Name=2;
CC CC      Isoid=060259-2; Sequence=VSP 005401;
CC CC      -!- TISSUE SPECIFICITY: Isoform 1 is predominantly expressed in the
CC CC      pancreas while isoform 2 is expressed in adult brain and
CC CC      hippocampus. Both forms are also found in fetal brain and
CC CC      placenta. Not detected in kidney, spleen, liver and lung.
CC CC      -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
DR DR      EMBL; AB009849; BAA28673.1; -
DR DR      EMBL; AB012761; BAA28676.1; -
DR DR      EMBL; AB010780; BAA88684.1; -
DR DR      EMBL; AB008390; BAA82665.1; -
DR DR      EMBL; AB008927; BAA82666.1; -
DR DR      EMBL; AF055982; AAD56050.1; -
DR DR      EMBL; AF095742; AAD25979.1; -
DR DR      EMBL; AF095743; AAD29574.1; -
DR DR      EMBL; AF243527; AAG33361.1; -
DR DR      EMBL; AC011473; AAG23254.1; -
DR DR      HSSP; O61955; INPM.
DR DR      MEROPS; S01.244; -
DR DR      Genew; HGNC:6369; KUK8.
DR DR      MIM; 605644; -
DR DR      GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR DR      GO; GO:0007399; P:neurogenesis; TAS.
DR DR      InterPro; IPR009003; Cys Ser trypsin.
DR DR      InterPro; IPR001254; Peptidase_S1.
DR DR      InterPro; IPR001314; Peptidase_S1A.
DR DR      Pfam; PF00089; trypsin; 1.
DR DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR DR      SMART; SM00020; Tryp_SPC; 1.
DR DR      PROSITE; PS00240; TRYPSIN_DOM; 1.
DR DR      PROSITE; PS00134; TRYPSIN_HIS; 1.
DR DR      PROSITE; PS00135; TRYPSIN_SER; 1.
DR DR      Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KW      Alternative splicing.
FT      SIGNAL 1 28 POTENTIAL.
FT      PROPEP 29 32 BY SIMILARITY.
FT      CHAIN 33 260 NEUROPSIN.
FT      ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      DISULFID 39 173 BY SIMILARITY.
FT      DISULFID 58 74 BY SIMILARITY.
FT      DISULFID 145 246 BY SIMILARITY.
FT      DISULFID 152 218 BY SIMILARITY.
FT      DISULFID 184 198 BY SIMILARITY.
FT      DISULFID 208 233 BY SIMILARITY.
FT      CARBOHYD 110 110 N-LINKED (GLCNAC...) (POTENTIAL).
FT      VARSPIC 23 23 A -> AACGSLDLTKLVAENLPVHLNPQWPSQSHCPRG
FT      WRSNLPRAA (in isoform 2).
FT      /FTID=VSP 005401.
SQ      SEQUENCE 260 AA; 28048 MW; EF439E5B8C3B660 CRC64;
Query Match 45.9%; Score 630.5; DB 1; Length 260;
Best Local Similarity 50.2%; Pred. No. 1.1e-44;
Matches 123; Conservative 24; Mismatches 93; Indels 5; Gaps 3;

```

```

QY 121 TSSVQPLPLPNDCATAGTECHVSGKGTNHRNPPDQLQCLNLSIVSHATCHGVYPGRI 180
DB 133 GSKVKPISLADHCTQPGQKQTVSGWGTVPSTPRENFPDTINCAEVKIFPKCKEDAYPGOI 192
QY 181 TSNMVCAGGVQDQACQDSDGGPLVCGGVLOGLVSWGSGVPGCGDGIQGVYTYICKYVDW 240
DB 193 TDGMVCAGSSKSGATCCQDSDGGPLVCDALQGITSWGS-DFGGRSDKFGVYTNICRYLDW 251
QY 241 IRMIM 245
DB 252 IKKII 256
RESULT 3
NRPN MOUSE
ID NRPN MOUSE STANDARD; PRT; 260 AA.
AC Q61955;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neopsin precursor (EC 3.4.21.-) (NP) (Kallikrein 8).
GN KUK8 OR PRSS19 OR NRPN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Hippocampus;
RX MEDLINE=95334817; PubMed=7623137;
RA Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,
RA Ito J., Nishino H., Aimoto S., Kiyama H., Shiosaka S.;
RT "Expression and activity-dependent changes of a novel limbic-serine
RT protease gene in the hippocampus.";
RL J. Neurosci. 15:5088-5097(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Yoshida S., Hirata A., Inoue N., Shiosaka S.;
RT "Cloning and assignment of mouse neopsin gene, Prss19 to chromosome
RT 7B4.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II; TISSUE=Mammary gland;
RX MEDLINE=22380257; PubMed=12477932;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF N-TERMINUS, AND CHARACTERIZATION.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=98225202; PubMed=9556608;
RA Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,
RA Shiosaka T., Midorikawa R., Kamachi T., Kawabe A., Shiosaka S.;
RT "Characterization of recombinant and brain neopsin, a

```

RT plasticity-related serine protease.";
 RL J. Biol. Chem. 273:11189-11196(1998).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 33-257.
 RC TISSUE=Hippocampus;
 RX MEDLINE=99134351; PubMed=9933620;
 RA Kishi T., Kato M., Shimizu T., Kato K., Matsumoto K., Yoshida S.,
 RA Shiozaki S., Hakoshima T.;
 RT "Crystal structure of neuropsin, a hippocampal protease involved in
 RT kindling epileptogenesis";
 RL J. Biol. Chem. 274:4220-4224(1999).
 CC -1- FUNCTION: Suggested to be involved in kindling epileptogenesis and
 CC hippocampal plasticity. Has a strong proteolytic activity against
 CC fibronectin.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.
 CC -1- ENZYME REGULATION: Strongly inhibited by diisopropyl
 CC fluorophosphate, leupeptin and (4-aminophenyl)methanesulfonyl 1-
 CC fluoride.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed specifically in the limbic system of
 CC mouse brain and is localized at highest concentration in pyramidal
 CC neurons of the hippocampal CA1-3 subfields.
 CC -1- MASS SPECTROMETRY: MW=26613; METHOD=MALDI; RANGE=29-260.
 CC -1- MASS SPECTROMETRY: MW=26229; METHOD=MALDI; RANGE=33-260.
 CC -1- SIMILARITY: Belongs to peptidase family S1. Kalikrein subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D30785; BAA06451.1; -;
 CC EMBL: AB032202; BAA92435.1; -;
 CC EMBL: BC055895; AAH55895.1; -;
 CC F1R; I56559; I56559.
 CC PDB; 1NPM; 23-NAR-99.
 CC MEROPS; S01.244; -;
 CC MGD; MGI:892018; Kik8.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR001254; Peptidase S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYP_SP; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW 3D-structure.
 KW SIGNAL 1 28 POTENTIAL.
 FT PROPEP 29 32
 FT CHAIN 33 260
 FT ACT_SITE 73 73 NEUROPSIN.
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 39 173
 FT DISULFID 58 74
 FT DISULFID 145 246
 FT DISULFID 152 218
 FT DISULFID 184 198
 FT DISULFID 208 233
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT STRAND 34 34
 FT STRAND 37 38
 FT TURN 41 42
 FT TURN 45 46
 FT STRAND 47 52
 FT TURN 53 54
 FT STRAND 55 64
 FT TURN 65 66

FT STRAND 67 70
 FT HELIX 72 74
 FT STRAND 80 83
 FT STRAND 87 87
 FT TURN 88 89
 FT STRAND 96 98
 FT STRAND 100 105
 FT TURN 107 108
 FT TURN 114 115
 FT TURN 118 119
 FT STRAND 122 126
 FT STRAND 140 141
 FT TURN 148 149
 FT STRAND 151 156
 FT STRAND 170 170
 FT STRAND 172 178
 FT HELIX 181 187
 FT TURN 189 191
 FT TURN 194 195
 FT STRAND 196 200
 FT TURN 202 203
 FT STRAND 206 206
 FT TURN 209 210
 FT TURN 212 213
 FT STRAND 215 218
 FT TURN 219 220
 FT STRAND 221 228
 FT STRAND 235 235
 FT TURN 236 237
 FT STRAND 238 238
 FT STRAND 240 244
 FT HELIX 245 256
 SQ SEQUENCE 260 AA; 28523 MW; BE5F6F6BE37CD60E CRC64;
 Query Match 45.3%; Score 622.5; DB 1; Length 260;
 Best Local Similarity 49.0%; Pred. No. 5.1e-44;
 Matches 121; Conservative 28; Mismatches 93; Indels 5; Gaps 3;
 QY 5 IFLLLCV---LGLSQAATPKIFNGTECGRNQSPQWQVGLFEGTSLRCGGVLIDHRWVLTAA 61
 DB 13 ILLELFWGAWAGLTRAQGSKILEGRECIPIHSPQWQAALFQGERLICGGVLGDRWVLTAA 72
 QY 62 HCSGSRYWVRLGHSLSQLDTEQIRHSGFSVTHPGVLGAS-TSHEHDLRLRLPLRVY 120
 DB 73 HCKKQKYSVRLGDSLSQSRDQPEQETQVAQSIQPCYNNSNPEDHSDIMLILQNSANL 132
 QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHRNPPFDLLQCLNLSIVSHATCHGVYPGRI 180
 DB 133 GDKVKPQLANLCPKVGQKCIISGWTVTSPQENFNTLNCAEVKIVSQNKCEAYPGKI 192
 QY 181 TSNMVCAGVPGODACQDGSGLVCGVLOGLVNWSVGPCCGQDGIPIGVYTYICKVDW 240
 DB 193 TEGMVCAGSSNGADTCQDGSGLVCDGMLQGITWSGS-DPCGKPKPGVYTKICRYTTW 251
 QY 241 IRMIMEN 247
 DB 252 IKKTMEN 258
 RESULT 4
 NRPN RAT
 ID NRPN RAT STANDARD; PRT; 260 AA.
 AC O88780;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuropsin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Brain serine
 DE protease 1).
 GN KIK8 OR PRSS19 OR NRPN OR BSP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;

```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=Fischer; TISSUE=Brain;
RX  MEDLINE=98389725; PubMed=9722524;
RA  Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathe R.;
RT  "Serine proteases in rodent hippocampus.";
RL  J. Biol. Chem. 273:23004-23011(1998)
CC  -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and
CC  hippocampal plasticity. Has a strong proteolytic activity against
CC  fibronectin [by similarity].
CC  -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.
CC  -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC  -!- TISSUE SPECIFICITY: Restricted to hippocampus.
CC  -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC  -----
CC  This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AJ005641; CAA06643.1; -.
DR  HSP; O61955; INPM.
DR  MEROPS; S01.244; -.
DR  InterPro; IPR009003; Cys_Ser_trypsin.
DR  InterPro; IPR001254; Peptidase_S1.
DR  InterPro; IPR001314; Peptidase_S1A.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  SMART; SM00020; TRYPSIN; 1.
DR  PROSITE; PS02040; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SSR; 1.
KW  Hydrolyase; Serine protease; Glycoprotein; Zymogen; signal.
FT  SIGNAL 1 28 POTENTIAL.
FT  PROPEP 29 32 BY SIMILARITY.
FT  CHAIN 33 260 NEUROPSIN.
FT  ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT  DISULFID 39 173 BY SIMILARITY.
FT  DISULFID 58 74 BY SIMILARITY.
FT  DISULFID 145 246 BY SIMILARITY.
FT  DISULFID 152 218 BY SIMILARITY.
FT  DISULFID 184 198 BY SIMILARITY.
FT  DISULFID 208 233 BY SIMILARITY.
FT  CARBOHYD 110 110 N-LINKED (GLCNAC...) (POTENTIAL).
SQ  SEQUENCE 260 AA; 58DF4F0602A0B7F5 CRC64;
Query Match 45.2%; Score 621.5; DB 1; Length 260;
Best Local Similarity 48.8%; Pred. No. 6.1e-44;
Matches 118; Conservative 32; Mismatches 90; Indels 3; Gaps 3;
QY 5 IFLLLCVGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLDHRVWLTAHC 63
DB 15 LFLMGAWAGLTRAQSKILEGQCKPHSPQWQTALEFQGERLVCGGVLDHRVWLTAHC 74
QY 64 SGRYVWRVLRGEHLSQLDWTQTRHSGFVTHPGYLGAS-TSHEHLRLRLRLPVRVTS 122
DB 75 KDKYSVRLGDHSLQKRDEPEQEIQVARSITQHPFCNFPNSDEHSHDMLRLQNSANLGD 134
QY 123 SVQPLPLNDCAATAGTECHVSGWGITHTHPNPDPDILLOCLNLISVSHATGCHVYVGRITS 182
DB 135 VKVPIELANLCPKVQKCIISGWTGVTSPQENFENTLCAEVKIYSQNKERAYPGKITE 194
QY 183 NMVCAAGVPGQDACQDGSGLVGLCGVGLQGLVSWGSGVPGQDGPVGVYTCIKYVDWIR 242
DB 195 GWCAGSSNGADTCQDGSGLVGLCGVGLQGLITWGS-DPCQKPKGVYTKICRYTNWIK 253
QY 243 MIM 245

```

```

Db 254 KTM 256
RESULT 5
KLIB HUMAN STANDARD; PRT; 250 AA.
AC Q9UBX7; O75837; Q9NS65;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Kallikrein 11 precursor (EC 3.4.21.-) (Hippostasin) (Trypsin-like
DE protease).
GN KLIK11 OR PRSS20 OR TLSP.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hippocampus;
RX MEDLINE=98438738; PubMed=9765601;
RY Yoshida S., Ianiguchi M., Suemoto T., Oka T., He X.P., Shiosaka S.;
RT "cDNA cloning and expression of a novel serine protease, TLSP.";
RL Biochim. Biophys. Acta 1399:225-228(1998).
[2]
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Hippocampus, and Prostate;
RX MEDLINE=20329229; PubMed=10872828;
RY Mitui S., Yamada T., Okui A., Kominami K., Uemura H., Yamaguchi N.;
RT "A novel isoform of a kallikrein-like protease, TLSP/hippostasin,
RT (PRSS20), is expressed in the human brain and prostate.";
RL Biochem. Biophys. Res. Commun. 272:205-211(2000).
[3]
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20130117; PubMed=10662548;
RY Yousef G.M., Scorrilas A., Diamandis E.P.;
RT "Genomic organization, mapping, tissue expression, and hormonal
RT regulation of trypsin-like serine protease (TLSP PRSS20), a new
RT member of the human kallikrein gene family.";
RL Genomics 63:88-96(2000).
[4]
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20510030; PubMed=11054574;
RY Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RY Moss P., Paepker B., Wang K.;
RA "Sequencing and expression analysis of the serine protease gene
RA cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
[5]
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RY Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RY Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RY Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,
RY Dangnanan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RY Andreise T., Trankhem M., Attix C., Amico-Keller G., Coefield J.,
RY Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RY Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RY Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[6]
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RY Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RY Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RY Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RY Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RY Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RY Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RY Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RY Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RY Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smail M.A.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Possible multifunctional protease. Efficiently cleaves
 CC bz-Phe-Arg-4-methylcoumaryl-7-amide, a kallikrein substrate, and
 CC weakly cleaves other substrates for kallikrein and trypsin.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9UBX7-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9UBX7-2; Sequence=VSP_005402;
 CC -!- TISSUE SPECIFICITY: Expressed in brain, skin and prostate. Isoform
 CC 1 is expressed preferentially in brain; isoform 2 in prostate.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB012917; BAA33404.1; AUT_INIT.
 CC EMBL; AB013730; BAA88713.1; -
 CC EMBL; AB014036; BAA96797.1; -
 CC EMBL; AF164623; AAD47815.1; -
 CC EMBL; AF243527; AAG33364.1; -
 CC EMBL; AC011473; AAG33257.1; -
 CC EMBL; BC022068; AAH22068.1; -
 CC HSSP; P00763; IDPO.
 CC MEROPS; S01.257; -
 CC Genew; HGNC:6359; KKL11.
 CC MiW; 60434; -
 CC GO; GO:0008236; F:serine-type peptidase activity; TAS.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen;
 KW Alternative splicing
 FT SIGNAL 1 18
 FT PROPEP 19 21
 FT CHAIN 22 250
 FT ACT_SITE 62 62
 FT ACT_SITE 110 110
 FT ACT_SITE 203 203
 FT DISULFID 28 163
 FT DISULFID 47 63
 FT DISULFID 135 237
 FT DISULFID 142 209
 FT DISULFID 174 188
 FT DISULFID 199 224
 FT CARBOHYD 99 99
 FT CARBOHYD 165 165
 FT CARBOHYD 181 181
 FT CARBOHYD 210 210
 FT VARSPIC 1 1
 M -> MQLRLWRDMKSSGRGLTAKEPGARSSPLQAM

FT (in isoform 2).
 FT /FTId=VSP_005402.
 SQ SEQUENCE 250 AA; 27466 MW; 192D910BBCDC7A56 CRC64;
 Query Match 45.0%; Score 618.5; DB 1; Length 250;
 Best Local Similarity 48.2%; Pred. No. 1e-43;
 Matches 120; Conservative 35; Mismatches 89; Indels 5; Gaps 4;
 QY 3 LSFLLLCVLGSLQAATPKIFNGTECGRNSQPWQGLFEGTSRLCGVLIIDHRWVLTAAH 62
 DB 4 LQILLALATGLVGGGT-RIIKGFECKPSQWQAALFEKTRLLCGATLIAPRWLTAAH 62
 QY 63 CSGRYVRLGHSLSQLDQTEQIRHSGFSVTHPGYLGA--STSHEDRLRLRLPLFVR 120
 DB 63 CLKPRIVHLGQNLQKEGCEQTRATSFPHGFNNSLPNKDHRNDMLVQWASPVSI 122
 QY 121 TSSVQPLPLPNDCAATAGTECHVSGWGITNHRNPPFDLLOCLNLSIVSHATCVPGRI 180
 DB 123 TWAVRPLTLSSRCVTAGTSLISGWSGTSPPQLRPLHTLRCAITITIEHQKCNAYPGNI 182
 QY 181 TSNMVCAG-GVPGQACQGDGGLVCGGLVGSVSGPCGQDGI PGVTVYICKYVD 239
 DB 183 TDTMVCASVQEGKQSGQSGGLVNCOSLQGIISWGQ-DPCAITRKPVGTVYICKYVD 241
 QY 240 WIRMENNN 248
 DB 242 WIQETMKN 250
 RESULT 6
 KLF HUMAN STANDARD; PRT; 256 AA.
 ID_KLKF_HUMAN Q9H2R3; Q9H2R4; Q9H2R6; Q9HBG9;
 AC Q9H2R5; Q15358; Q9H2R3; Q9H2R4; Q9H2R6; Q9HBG9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Kallikrein 15 precursor (BC 3.4.21.-) (ACO protease).
 GN KKL15.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RX PubMed=11010966;
 RA Yousef G.M., Scorilas A., Jung K., Ashworth L.K., Diamandis E.P.;
 RT "Molecular cloning of the human kallikrein 15 gene (KLK15). Up-
 RT regulation in prostate cancer.";
 RL J. Biol. Chem. 276:53-61(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20510030; PubMed=11054574;
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Paepker B., Wang X.;
 RT "Sequencing and expression analysis of the serine protease gene
 RT cluster located in chromosome 19q13 region.";
 RL Gene 257:119-130(2000).
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 RA MEDLINE=94289486; PubMed=8018728;
 RA Dihanich M.E., Spiess M.;
 RT "A novel serine proteinase-like sequence from human brain.";
 RL Biochim. Biophys. Acta 1218:225-228(1994).
 CC -!- FUNCTION: Protease whose physiological substrate is not yet known.
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1;
 CC IsoId=Q9H2R5-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9H2R5-2; Sequence=VSP_005405;
 CC Name=3;

```
CC      IsoId=Q9H2R5-3; Sequence=VSP_005406; VSP_005407;
CC      Name=4;
CC      IsoId=Q9H2R5-4; Sequence=VSP_005404;
CC      TISSUE SPECIFICITY: Highest expression in the thyroid gland. Also
CC      expressed in the prostate, salivary, and adrenal glands and in the
CC      colon testis and kidney.
CC      -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AF242195; AAG09469.1; -
CC      EMBL; AF242195; AAG09470.1; -
CC      EMBL; AF242195; AAG09471.1; -
CC      EMBL; AF242195; AAG09472.1; -
CC      EMBL; AF243527; AAG33354.1; -
CC      EMBL; X75363; CAA53145.1; ALT_SEQ.
CC      HSP; P00763; IDPO.
CC      MEROPS; S01.081; -.
CC      GO; GO:000576; C:extracellular; NAS.
CC      GO; GO:0004252; P:serine-type endopeptidase activity; NAS.
CC      GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
CC      InterPro; IPR009003; Cys_Ser_trypsin.
CC      InterPro; IPR001254; Peptidase_S1.
CC      InterPro; IPR001314; Peptidase_S1A.
CC      Pfam; PF00089; trypsin; 1.
CC      PRINTS; PR00722; CHYMOTRYPSIN.
CC      PROSITE; PS50240; TRYPSIN_DOM; 1.
CC      PROSITE; PS00134; TRYPSIN_HIS; 1.
CC      PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
CC      Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen;
KW      Alternative splicing.
FT      SIGNAL      1 16      POTENTIAL.
FT      PROPEP      17 21      ACTIVATION PEPTIDE (POTENTIAL).
FT      CHAIN      22 256      KALLIKREIN 15.
FT      ACT_SITE    62 62      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE    106 106     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE    209 209     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      CARBOHYD     171 171     N-LINKED (GLCNAC...) (POTENTIAL).
FT      CARBOHYD     232 232     N-LINKED (GLCNAC...) (POTENTIAL).
FT      VARSPLIC    122 206     Missing (in isoform 4).
FT      VARSPLIC    122 256     Missing (in isoform 2).
FT      VARSPLIC    161 161     V -> G (in isoform 3).
FT      VARSPLIC    162 256     /FTid=VSP_005406.
FT      VARSPLIC    147 160     Missing (in isoform 3).
FT      CONFLICT    147 160     SHNEPGTAGSPRSQ -> PLSSP (IN REF. 2).
FT      SEQUENCE    256 AA; 28087 MW; 5BEBF8D602786B5 CRC64;
Query Match      43.6%; Score 599; DB 1; Length 256;
Best Local Similarity 47.1%; Pred. No. 4.2e-42;
Matches 121; Conservative 32; Mismatches 88; Indels 16; Gaps 6;
QY      5 IFLLVLVLGSLQAAT--PKIFNGTCGRNSQWQVGLFEGTSLRCGGVLIDHRVLTAAH 62
DB      3 LLTLTSLASTAAQDGKLLGDECAHPQWQVALYERCFNCGASLLSPHWLSAAH 62
QY      63 CSGRYVWRLGEHSLQDWTQIRHSGFVTHPGYLGASTSHEDHRLRLRLPVRVTS 122
DB      63 CSQRFMRVRLGEHNLRLKRDGPEQLRTSRVIPHRY--EARSHRNDLMLRLVQPARLNP 120
QY      123 SVQPLPLPNDCATAGTCHVSGMGIYTH-----PRN--PFDDLQCLNLSVSHATC 172
DB      121 QVRPAVLFPTRCPHPGEACVSGWGLVSHNEPGTAGSPRSQVSLPDTLHCANISLIISDTC 180
QY      173 HGVPYGRITSNMVCAGGV-PQDACQGDGGGLVCGGVLQGLVSGWGVCGQGDGIPGVY 231
```

```
DB      181 DKSPYGRITNTMVCAGAEGRGAECEGSGGLVCGGVLQGLVSGWGV-PCDNTTKPGVY 239
QY      232 TYICKYVDWIRIMMNN 248
DB      240 TKVCHYLEWIRETKRN 256
RESULT 7
TRY3_CHICK
ID TRY3_CHICK STANDARD; PRT; 248 AA.
AC Q90629;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trypsin II-P29 precursor (EC 3.4.21.4).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=95251611; PubMed=7733885;
RA Wang K., Gan L., Lee I., Hood L.E.;
RT Isolation and characterization of the chicken trypsinogen gene
RL Biochem. J. 307:471-479 (1995).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: High levels are seen in the pancreas while
CC lower levels are found in the liver, spleen and thymus.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U15157; AAA79914.1; -.
CC PIR; S55066; S55066.
CC HSP; P00763; IDPO.
CC MEROPS; S01.151; -.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW Calcium-binding; Signal; Multigene family.
FT      SIGNAL      1 16      BY SIMILARITY.
FT      PROPEP      17 25      ACTIVATION PEPTIDE (BY SIMILARITY).
FT      CHAIN      26 248      TRYPSIN II-P29.
FT      ACT_SITE    65 65      CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      METAL       77 77      CALCIUM (BY SIMILARITY).
FT      METAL       79 79      CALCIUM (VIA CARBONYL OXYGEN)
FT (BY SIMILARITY).
FT      METAL      82 82      CALCIUM (VIA CARBONYL OXYGEN)
FT (BY SIMILARITY).
FT      METAL      87 87      CALCIUM (BY SIMILARITY).
FT      ACT_SITE    109 109     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE    202 202     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      DISULFID     32 162     BY SIMILARITY.
FT      DISULFID     50 66      BY SIMILARITY.
```

FT DISULFID 134 235 BY SIMILARITY.
 FT DISULFID 141 208 BY SIMILARITY.
 FT DISULFID 173 187 BY SIMILARITY.
 FT DISULFID 198 222 BY SIMILARITY.
 FT SITE 196 196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 248 AA; 26622 MW; E5E16B07622B588E CRC64;

Query Match 41.4%; Score 569.5; DB 1; Length 248;
 Best Local Similarity 47.5%; Pred. No. 1e-39;
 Matches 116; Conservative 33; Mismatches 78; Indels 11; Gaps 6;

QY 5 IFLLLCVLGLS-----QAATPKIFNGTECGRNQPMQVGLFEGTSURCGGVLDHRWLT 59
 DB 4 LFLILSLGLAFAVPPGADDDKIVGGVTCPEHSVYQVSLNSGVHF--CGGLNSINQWLS 62

QY 60 AAMGSGRYWRGLGHSLSQDWTQIRHSGFSVTHGVLGASTSHEHDLRLRLR 119
 DB 63 AAHCYSRIQVRIGEXYIDVQEDSEVRSSVIRHPKY--SSITLNDIMLKASAVE 120

QY 120 VTSVQPLPNDPCATAGTECHVSGWGTINHPNPFPLDQLCLMLSTIVSHATCHGVY 179
 DB 121 YSADIQIPALPSSCAKAGTECLISGNGYLSNGYNYPELLOCLNAPILSDQCEAYPGD 180

QY 180 ITSNMVCAGGVP--GQDACQSGGFLVCGGVGLQGLVSGVSGPGQDGIPIGVYTI 238
 DB 181 ITSNMVCGLGEGKQSGDGGPVGNGELQGVSG-IG-CALXGYPGVYTKVNYV 238

QY 239 DWIR 242
 DB 239 DWIQ 242

RESULT 8
 ID KLKE_HUMAN STANDARD; PRT; 251 AA.
 AC Q9P0G3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Kallikrein 14 precursor (EC 3.4.21.-) (Kallikrein-like protein 6)
 DE (KLU-6).
 GN KLU14 OR KLU6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yousef G.M., Diamandis E.P.;
 RT "Molecular characterization, mapping, and tissue expression of KLU6,
 a hormonally regulated kallikrein-like gene."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=21250997; PubMed=11352573;
 RA Hooper J.D., Bui L.T., Rae F.K., Harvey T.J., Myers S.A.,
 RA Ashworth L.K., Clements J.A.;
 RT "Identification and characterization of KLU14, a novel kallikrein
 serine protease gene located on human chromosome 19q13.4 and expressed
 in prostate and skeletal muscle."
 RL Genomics 73:117-122(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Barnes J.,
 RA Danganan L., Eriker A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.;
 RT "Sequence analysis of chromosome 19q13.4."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20545474; PubMed=10969073;
 RA Harvey T.J., Hooper J.D., Myers S.A., Stephenson S.A., Ashworth L.K.,
 RA Clements J.A.;
 RT "Tissue-specific expression patterns and fine mapping of the human
 kallikrein (KLU) locus on proximal 19q13.4."
 RL J. Biol. Chem. 275:37397-37406(2000).
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -!- TISSUE SPECIFICITY: High expression in brain, bone marrow and
 CC fetal liver. Also expressed in liver, pancreas, fetal spleen,
 CC prostate and skeletal muscle.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF161221; AAD50773.2; -;
 CC EMBL; AF283669; AAK48523.1; -;
 CC EMBL; AF283670; AAK48524.1; -;
 CC EMBL; AC011473; AAC23260.1; -;
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.029; -;
 CC Genew; HGNC:6362; KLU14.
 CC MIM; 606135; -;
 CC GO; GO:0005576; C:extracellular; NAS.
 CC GO; GO:0004252; P:serine-type endopeptidase activity; NAS.
 CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_SPc; 1.
 CC PROSITE; PS0240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Signal; zymogen.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 25 251 KALLIKREIN 14.
 FT ACT_SITE 67 67 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 111 111 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 31 164 BY SIMILARITY.
 FT DISULFID 52 68 BY SIMILARITY.
 FT DISULFID 143 210 BY SIMILARITY.
 FT DISULFID 175 189 BY SIMILARITY.
 FT DISULFID 200 225 BY SIMILARITY.
 SQ SEQUENCE 251 AA; 27452 MW; 9087953BAPAF7ED25 CRC64;

Query Match 41.4%; Score 569.5; DB 1; Length 251;
 Best Local Similarity 46.9%; Pred. No. 1e-39;
 Matches 119; Conservative 33; Mismatches 87; Indels 15; Gaps 6;

QY 5 IFLLLCVLGLS-----SQAATPKIFNGTECGRNQPMQVGLFEGTSUR--CGGVLDHR 55
 DB 1 MFLLLTALQVLAIAMTQSQDENKIIGGHTCTRSSQPMQVQALLAGRRFLCGGALLSGQ 60

QY 56 VWLTAHSGSYWRVRLGHSLSQDWTQIRHSGFSVTHGVLGASTSHEHDLRLRLR 115
 DB 61 WVTTAHCGRPILQVALGKHLRERMEATQVLRVVRQVTHPNY--NSRTHDNDMLLQLQ 118

QY 116 LPVRTSSVQPLPLPNDPCATAGTECHVSGWGTINHPNPFPLDQLCLMLSTIVSHATCHGV 175
 DB 119 QPARIGRAVRPEVTQACASPGTSCVSGWGTISSPIARYPASLQCVNINISPDVCQKA 178

QY 176 YPGRTITSNMVCAGGVP--GQDACQSGGFLVCGGVGLQGLVSGVSGPGQDGIPIGVY 233

Db 179 YPRITFGMCA-GVPGGKDSQGGSLVCRGLQGLVSMG-MERCALPGPGVYTN 236
Qy 234 ICKYVDWIRMN 247
Db 237 LCKYRSWIEETMRD 250

DR MEROPS: S01.307; -
DR Genew; HGNC:6370; KLK9.
DR MIM; 605504; -
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; CysSer trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Signal.
FT SIGNAL 1 15
FT CHAIN 16 250 KALLIKREIN 9.
FT ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 111 111 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 29 164 BY SIMILARITY.
FT DISULFID 48 64 BY SIMILARITY.
FT DISULFID 136 238 BY SIMILARITY.
FT DISULFID 143 210 BY SIMILARITY.
FT DISULFID 175 189 BY SIMILARITY.
FT DISULFID 200 225 BY SIMILARITY.
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 250 AA; 27512 MW; F2785245B063E98B CRC64;

Query Match 41.4%; Score 568.5; DB 1; Length 250;
Best Local Similarity 46.7%; Pred. No. 1.3e-39;
Matches 119; Conservative 32; Mismatches 89; Indels 15; Gaps 7;
Qy 1 MGLSIFILLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGVLIDHRWLTA 60
Db 3 LGLICALLSLLAGHWADT-RAIGAECRFPNSQPWQAGLFLHTRFCGATLSDRWLTA 61
Qy 61 AHCSGSYVWVRLGHSLSQLDWTEQIRHSGFSVT----HPGY---LGASTSHEHDLRLR 113
Db 62 AHCRKPYLWVRLGHEHLWKWEGPEQL-----FRVTDFFPHGFGNKDLSAN-DHNDIMLR 116
Qy 114 LRLPVRTSSVQPLPLPNDCATAGTECHVSGWGITNHPNPFDDLLQCLNLSIVSHATCH 173
Db 117 LPROARLSPAVQPLNLSOTCVSPGMQLISGWAVSPKALFPVTLQCANISLENKLCH 176
Qy 174 GVPGRITSNMVCAG-GVPGQDACQGDGSGPLVCGVLQGLVSMGSGVPGCGQDGIPGVY 232
Db 177 WAYPGHISDSMLCAGLWEGGRSGCGDGGPLVCGVNTLAGVVS-GGAEPCSRPRPAVVT 235
Qy 233 YICKYVDWIRMN 247
Db 236 SVCHYLDWIQIEMEN 250

RESULT 10
KLK9 HUMAN
ID KLK9 HUMAN STANDARD; PRT; 250 AA.
AC Q9UKR3; Q9Y433;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kallikrein 9 precursor (EC 3.4.21.-) (Kallikrein-like protein 3) (KLK-L3).
GN KLK9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=20118156; PubMed=10652563;
RA Yousef G.M., Iuo L.-Y., Diamandis E.P.;
RT "Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4.";
RL Anticancer Res. 19:2843-2852(1999).
RN [2]
RP SEQUENCE FROM N.A.
MEDLINE=20247258; PubMed=10783266;
RA Yousef G.M., Diamandis E.P.;
RT "The expanded human kallikrein gene family: locus characterization and molecular cloning of a new member, KLK-13.";
RL Genomics 65:184-194(2000).
RN [3]
RP SEQUENCE FROM N.A.
MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepel B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J., Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Andreise T., Trankheim M., Attix C., Amico-Keller G., Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmuller B., Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4.";
RL Submitted (OCT-2000) to the EMBL/Genbank/DDay databases.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: Skin, thymus, trachea, cerebellum and spinal cord.
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF135026; AAD26427.2; -
CC EMBL; AF243527; AAG33362.1; -
CC EMBL; AC011473; AAG23255.1; -
CC HSSP; P00763; 1DPO.

RX MEDLINE=20229789; PubMed=10766816;
RA Yousef G.M., Chang A., Diamandis E.P.;
RT "Identification and characterization of KLK-L4, a new kallikrein-like
RL gene that appears to be down-regulated in breast cancer tissues.";
RN J. Biol. Chem. 275:11891-11898(2000).
[2]
SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,
RA Danganan L., Eriar A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of chromosome 19q13.4";
RN Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE OF 1-180 FROM N.A.
RC TISSUE=Uterus;
RA Ansoez W., Winkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in prostate, breast, testis and
CC salivary gland.
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF135024; AAD26425.2; -;
CC EMBL; AC011473; AAG23259.1; -;
CC EMBL; AL050220; CAB43320.1; ALT_INIT.
CC HSP; P00763; IDPO.
CC MEROPS; S01.306; -;
CC Genew; HGNC:6361; KLU13.
CC MIM; 605505; -;
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYSPIN_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 277 KALLIKREIN 13.
FT ACT_SITE 76 76 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 124 124 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 178 BY SIMILARITY.
FT DISULFID 61 77 BY SIMILARITY.
FT DISULFID 157 224 BY SIMILARITY.
FT DISULFID 189 203 BY SIMILARITY.
FT DISULFID 214 239 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 170 180 VNPVKTLQCAN -> GMHPRWPEAP (IN REF. 3).
FT SEQUENCE 277 AA; 30570 MW; BA8A9B8DFB5D542 CRC64;
Query Match 41.3%; Score 567; DB 1; Length 277;
Best Local Similarity 46.8%; Pred. No. 1.9e-39;
Matches 116; Conservative 36; Mismatches 80; Indels 16; Gaps 6;

QY 13 GLSQAAATPKIFN-----GTEGRNSQWQVGLFEGTSIRCGVLIIDHRWVLTAAH 62
DB 18 GVSQESS-KVLNTNGTSGFLPGGYTCFPHSQWQAAALLVQGRLLCGVLIHVPKWVLTAAH 76
QY 63 CSGSRVWVRLGHSLSOLDWTEQIRHSFGSVTHPGVILGAST--SHEDLRLRLRLPVRV 120
DB 77 CLKEGKVLGRALGRVAGQVREVVISIHPPEYRSPFTHLNHDDHIMLLEQSPVQL 136
QY 121 TSSVQELPLP-NDCAATAGTCHVSGWGI TNHPRNPDPDLLOCLNLSIVSHATCHGVYGR 179
DB 137 TGYIQLPLSHNNRLPGTTCRVSGWGTTSPOVNYPKTLQCANIQLRSDDECRQVYPGK 196
QY 180 ITSNMVCAGVGP-QDACQDGSGLVCGVLOGLVSGVSGVPCGOGGIPGVVYTIKVV 238
DB 197 ITDNMLCAGTKEGKDCSCBDSGGLVNCNTLYGVSWGDF-PCGQDPRPGVTRVSRVY 255
QY 239 DWIRMIWR 246
DB 256 LWIREIR 263
RESULT 11
TRY2_CHICK
ID TRY2_CHICK STANDARD; PRT; 248 AA.
AC Q90628;1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Trypsin 1-P38 precursor (EC 3.4.21.4).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=95251611; PubMed=7733885;
RA Wang K., Gan L., Lee I., Hood L.E.;
RT "Isolation and characterization of the chicken trypsinogen gene
RL family.";
RL Biochem. J. 307:471-479(1995).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: High levels are seen in the pancreas while
CC lower levels are found in the liver, spleen and thymus.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U15156; AAA79913.1; -;
CC PIR; S55067; S55067.
CC HSP; P00763; IDPO.
CC MEROPS; S01.258; -;
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYSPIN_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen;
KW Calcium-binding; Signal; Multigene family.

RX MEDLINE=96320486; PubMed=8764136;
 RA Liu X.-L., Wazer D.E., Watanabe K., Band V.;
 RT "Identification of a novel serine protease-like gene, the expression
 RL of which is down-regulated during breast cancer progression.";
 RN Cancer Res. 56:3371-3379(1996).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99321170; PubMed=9647736;
 RX Luo L.-Y., Herbrick J.A., Scherer S.W., Beatty B., Squire J.,
 RA Diamandis E.P.;
 RT "Structural characterization and mapping of the normal epithelial
 RL cell-specific 1 gene";
 RN Biochem. Biophys. Res. Commun. 247:580-586(1998).
 [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20510030; PubMed=11054574;
 RX Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Paepker B., Wang K.;
 RT "Sequencing and expression analysis of the serine protease gene
 RL cluster located in chromosome 19q13 region.";
 RN Gene 257:119-130(2000).
 [4]
 RN SEQUENCE FROM N.A.
 RP Lanerding J.E., McCready P.M., Skowronski E., Viewanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,
 RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Cosfield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.;
 RT "Sequence analysis of chromosome 19q13.4";
 RN Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 [5]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Uterus;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.N., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [6]
 RN CHARACTERIZATION.
 RP MEDLINE=99025848; PubMed=9809976;
 RX Goyal J., Smith K.M., Cowan J.M., Wazer D.E., Lee S.W., Band V.;
 RA "The role for NES1 serine protease as a novel tumor suppressor";
 RT Cancer Res. 58:4782-4786(1998).
 CC -!- FUNCTION: Has a tumor-suppressor role for NES1 in breast and
 CC prostate cancer.
 CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -!- TISSUE SPECIFICITY: Expressed in breast, ovary and prostate.
 CC -!- DEVELOPMENTAL STAGE: Down-regulated during breast cancer
 CC progression.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF024605; AAB81602.1; -
 DR EMBL; AF055481; AAC14266.1; -
 DR EMBL; AF243527; AAG33363.1; -
 DR EMBL; AC011473; AAG23256.1; -
 DR EMBL; BC002710; AAH02710.1; -
 DR HSSP; P00763; 1DPO.
 DR MEROPS; S01.246; -
 DR Genew; HGNC:6358; KKL10.
 DR MIM; 602673; -
 DR GO; GO:0005576; C:extracellular; TAS.
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM0020; TRYP_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS0134; TRYPSIN_HIS; 1.
 DR PROSITE; PS0135; TRYPSIN_SER; FALSE NEG.
 KW Hydroxylase; Serine protease; Glycoprotein; Signal; Anti-oncogene.
 FT SIGNAL 1 30
 FT CHAIN 31 276
 FT KALLIKREIN 10.
 FT ACT_SITE 86 86
 FT ACT_SITE 137 137
 FT ACT_SITE 229 229
 FT DISULFID 52 162
 FT DISULFID 71 87
 FT DISULFID ? 263
 FT DISULFID 169 235
 FT DISULFID 201 215
 FT DISULFID 225 250
 FT CARBOHYD 39 39
 FT CONFLICT 50 50
 FT CONFLICT 149 149
 FT CONFLICT P -> L (IN REF. 3 AND 4).
 SQ SEQUENCE 276 AA; 30138 MW; 82A2507379BAB313 CRC64;
 Query Match 41.0%; Score 564; DB 1; Length 276;
 Best Local Similarity 45.9%; Pred No. 3.3e-39;
 Matches 118; Conservative 29; Mismatches 92; Indels 18; Gaps 5;
 QY 7 LLLCVLGLSQA-----TPKIFNGTECRNSQPWQVGLFEGTSLRCGGVLIDHRWV 57
 DB 23 LLMAQLWAAEAALLPQNDRLDPEAY-GAPCARGSQPWQVSLFNLGLSPHCAGLVLDQSWV 81
 QY 58 LTAACSGSRVWVLGEHSLQSLDWTETQIRHSGSVTHPGVILGAS-----TSHEDLRL 111
 DB 82 LTAACGCKPLWARVGDHLLQLLQ-GEQLRRTTSVWHPKYHQGGGPIPLRRTDDEHML 140
 QY 112 LRLRLPVRVTSVQPLPLPNDCATAGTECHVSGWGITNHPNPFPLQLCLNLSTVSHAT 171
 DB 141 LKLAPVPGPRVRAQLPYRCAQPGDCCQVAGGTTAARVKYKNGKLTCSITILSPKE 200
 QY 172 CHGVYPGRITSNMVCGAGVPGDQACQSGSGPLVCGVQLGVLVSGVSGPGQDGPV 231
 DB 201 CEVFPYGVVNNMICAGLDRGQDPCQSDSGGGLVCDLQGLISWG-VYPCGSAQHPAVY 259
 QY 232 TYICKYVDMIRMINN 248
 DB 260 TOICKYMSWINKVRSN 276
 RESULT 15
 TRY2_XENLA
 ID TRY2_XENLA STANDARD; PRT; 244 AA.
 AC P70059;
 DT 01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Trypsin precursor (EC 3.4.21.4)
Xenopus laevis (African clawed frog).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
[1]
SEQUENCE FROM N.A.
Wang K., Lytle L., Gan L., Hood L.E.;
Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
-!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
-!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
-!- SUBCELLULAR LOCATION: Extracellular.
-!- SIMILARITY: Belongs to peptidase family S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U72330; AAB17274.1; -.
HSP; P00763; IDPO.
DR DR MEROPS; S01.258; -.
DR DR InterPro; IPR009003; Cys Ser trypsin.
DR DR InterPro; IPR001254; Peptidase S1.
DR DR InterPro; IPR001314; Peptidase_S1A.
DR DR Pfam; PF00089; trypsin; 1.
DR DR PRINTS; PR00722; CHYMOTRYPSIN.
DR DR SMART; SM00020; TRYPSIN_SPC; 1.
DR DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Zymogen; Calcium-binding;
KW Signal; Multigene family.
FT SIGNAL 1 15 BY SIMILARITY.
FT PROPEP 16 21 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 22 244 TRYPSIN.
FT ACT_SITE 61 61 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT METAL 73 73 CALCIUM (BY SIMILARITY).
FT METAL 75 75 CALCIUM (VIA CARBONYL OXYGEN)
FT METAL 83 83 (BY SIMILARITY).
FT METAL 83 83 CALCIUM (BY SIMILARITY).
FT ACT_SITE 105 105 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 198 198 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 28 158 BY SIMILARITY.
FT DISULFID 46 62 BY SIMILARITY.
FT DISULFID 130 231 BY SIMILARITY.
FT DISULFID 137 204 BY SIMILARITY.
FT DISULFID 169 183 BY SIMILARITY.
FT DISULFID 194 218 BY SIMILARITY.
FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
FT SEQUENCE 244 AA; 26079 MW; C63F29CB3300B323 CRC64;
Query Match 41.0%; Score 563; DB 1; Length 244;
Best Local Similarity 45.0%; Pred. No. 3.5e-39;
Matches 112; Conservative 46; Mismatches 83; Indels 8; Gaps 5;

QY 3 LSIFLLCVLGLSQA--ATPKIFNGTEGRNSQPMQVGLFEGTSIRCGGVLIDHRWLTA 60
DB 1 MKFLVLLVLLGAFAFEDDDDKIVGGFTCAKNAPYQVSLNAGYHP-CGGLINSQWVISA 59

QY 61 AHCSGSRYWVRIGBSLSQLDWTQTQIRHSGFSTVTHPGYLGASTSHEHDLRLRLRFPVR 120
DB 60 AHCKSRIQVRLGEHNIALNEGTFQIDSQVKIKHPNY--NSRNLNDNDIMLKLTSTTARL 117

QY 121 TSSVQPLPLPNDCTAGTECHVSGNGITNHPNRPFPDLLOCLINLSIVSHATCHGVPGRI 180
DB 118 SANIQSVPLPSACASAGTNCILISGWTNLSSTGTYNPDLLQCLINAPILTDSQCSNYPGEI 177

QY 181 TSNMYCAGGVP-GQACQGDSDGGPLVCGGVLOGLVSWGSVGPCCGGDGPVGYTYICKYVD 239
DB 178 TKNMFCAGFLAGGKDSQGDGGPVCNGLQGVVSWGY--GCAQRNYEPGVYTKVCNFVT 235
QY 240 WURMIMRNN 248
DB 236 WIQSTISSN 244

Search completed: June 16, 2004, 20:33:11
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 01:43:09 ; Search time 4624 Seconds
(without alignments)
10226.481 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcagtcaccccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1091	100.0	1091	6	AX697125	Sequence
2	1091	100.0	1091	9	AY358524	Homo sapi
3	510.8	46.8	811	9	BSM808731	Sequence
4	506.4	46.4	840	6	AR337568	Human ser
5	506.4	46.4	840	6	BD130047	Sequence
6	493.2	45.2	619	6	AR263961	Human ser
7	373.2	34.2	9120	9	AF164623	Homo sapi
8	373.2	34.2	11820	9	AF135025	Homo sapi
9	373.2	34.2	132323	9	AC011473	Homo sapi
10	373.2	34.2	230000	9	AF243527	Homo sapi
11	368.4	33.8	174724	2	AC140096	Pan trogl
12	368.4	33.8	200792	2	AC130782	Pan trogl
13	361.2	33.1	142334	2	AC073185	Homo sapi
14	297	27.2	405	6	AX070938	Sequence
15	294	26.9	176647	2	AC130188	Papio anu
16	219.4	20.1	1146	6	BD193094	207 human
17	217.8	20.0	1166	6	AR152173	Sequence
18	217.6	19.9	1106	6	AX016287	Sequence
19	217.6	19.9	1106	6	BD137019	Human CAS
20	217.6	19.9	1158	6	AX016289	Sequence
21	217.6	19.9	1158	6	BD137020	Human CAS
22	217.6	19.9	1181	9	AB013730	Homo sapi
23	217.6	19.9	1186	9	AB012917	Homo sapi
24	217.6	19.9	1192	6	AR152174	Sequence
25	217.6	19.9	1204	6	AX358932	Sequence
26	217.6	19.9	1204	6	AX362425	Sequence
27	217.6	19.9	1204	6	AX454622	Sequence
28	217.6	19.9	1204	6	AX464372	Sequence
29	217.6	19.9	1204	6	AX491100	Sequence
30	217.6	19.9	1204	6	AX697101	Sequence
31	217.6	19.9	1204	9	AY359014	Homo sapi
32	217.6	19.9	1213	9	BC022068	Homo sapi
33	217.6	19.9	1301	6	BD091587	Novel ser
34	217.6	19.9	1301	9	AB041036	Homo sapi
35	217.6	19.9	1314	6	AR098430	Sequence
36	217.6	19.9	1314	6	BD130920	Serine pr
37	216	19.8	1191	6	BD139483	Extended
38	210	19.2	1335	6	BD232165	A novel m
39	206.8	19.0	833	6	AR060847	Sequence
40	206.8	19.0	833	6	BD082136	Novel pro
41	200.4	18.4	1213	10	AB016226	Mus muscu
42	200.4	18.4	1256	10	AB016227	Mus muscu
43	200.4	18.4	1323	6	BD091588	Novel ser
44	199.4	18.3	1052	6	AR219287	Sequence
45	198	18.1	1283	9	AF283670	Homo sapi

ALIGNMENTS

RESULT 1
AX697125
LOCUS AX697125
DEFINITION Sequence 193 from Patent WO0078961.
ACCESSION AX697125
VERSION AX697125.1 GI:29498087
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ferrara, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, L.,
Eaton, D.L., Gao, W.Q., Pan, J., Botstein, D., Fong, S., Goddard, A.,
Godowski, P.J., Gurney, A.L., Smith, V., Tumas, D., Wood, W.I.,


```
Query Match      100.0%; Score 1091; DB 9; Length 1091;
Best Local Similarity 100.0%; Pred. No. 8.1e-238; Indels 0; Gaps 0;
Matches 1091; Conservative 0; Mismatches 0;

QY 1 CAAGCAGGTCTATCCCTTGGTGTGACCTTCAAAGAGAGAGAGAGAGAGAGTGGGGGC 60
DB 1 CAAGCAGGTCTATCCCTTGGTGTGACCTTCAAAGAGAGAGAGAGAGAGAGTGGGGGC 60

QY 61 ACAGGAAAGGTGTACCTCTGAGATCCCTTTTCCCGCAGACTTTGGAAGTGCACACC 120
DB 61 ACAGGAAAGGTGTACCTCTGAGATCCCTTTTCCCGCAGACTTTGGAAGTGCACACC 120

QY 121 ATGGGGCTCAGCATCTTTTGTCTGTGTCTTGGGCTCAGCCAGGACACACCG 180
DB 121 ATGGGGCTCAGCATCTTTTGTCTGTGTCTTGGGCTCAGCCAGGACACACCG 180

QY 181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAATCTACAGCCGTGGAGTGGGGTGT 240
DB 181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAATCTACAGCCGTGGAGTGGGGTGT 240

QY 241 GAGGCAACAGCTCGCTGCGGTGGGGGTGTCCTATTGACCAAGTGGGTCTCACAGG 300
DB 241 GAGGCAACAGCTCGCTGCGGTGGGGGTGTCCTATTGACCAAGTGGGTCTCACAGG 300

QY 301 GCTCACTGACGGGACGAGTACTGGGTGCGCTGGGGGAACACAGCCCTCAGCCAGCTC 360
DB 301 GCTCACTGACGGGACGAGTACTGGGTGCGCTGGGGGAACACAGCCCTCAGCCAGCTC 360

QY 361 GACTGACCGAGAGATCCGGACAGCGGTCTCTGTGACCGATCCCGGTCTACCTGGGA 420
DB 361 GACTGACCGAGAGATCCGGACAGCGGTCTCTGTGACCGATCCCGGTCTACCTGGGA 420

QY 421 GCTCGAAGCAGCAGCAGCAGCAGCTCGGCTGCTGCGCTGCGCTGCGCTGCGCTG 480
DB 421 GCTCGAAGCAGCAGCAGCAGCAGCTCGGCTGCTGCGCTGCGCTGCGCTGCGCTG 480

QY 481 ACCAGCAGGCTTCAACCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 540
DB 481 ACCAGCAGGCTTCAACCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 540

QY 541 CAGCTCTCAGGTGGGGATACCAACACCAACCGAAACCAATCCCGGATCTGCTCCAG 600
DB 541 CAGCTCTCAGGTGGGGATACCAACACCAACCGAAACCAATCCCGGATCTGCTCCAG 600

QY 601 TGCTCAACTCTCCATCTGCTCCATGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 660
DB 601 TGCTCAACTCTCCATCTGCTCCATGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 660

QY 661 ACAGCAACATGTTGTGTCAGCGCGGTCCCGGGCAGGATCCCTGCGAGGTGATTTCT 720
DB 661 ACAGCAACATGTTGTGTCAGCGCGGTCCCGGGCAGGATCCCTGCGAGGTGATTTCT 720

QY 721 GGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAAGGTCTGTTGTTGTTGTTGTTGTTG 780
DB 721 GGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAAGGTCTGTTGTTGTTGTTGTTGTTG 780

QY 781 CCTGTGGAACAATGTCATCCCTGAGTCTACACCTATATTTGCAAGTATGGAAGTCTG 840
DB 781 CCTGTGGAACAATGTCATCCCTGAGTCTACACCTATATTTGCAAGTATGGAAGTCTG 840

QY 841 ATCCGATGATCATGAGGAACAATGACCTGTTTTCCTCCACCTCCACCCCTTAA 900
DB 841 ATCCGATGATCATGAGGAACAATGACCTGTTTTCCTCCACCTCCACCCCTTAA 900

QY 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACAAATCTCTCTCCATCACTTCCCTAGTCC 960
DB 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACAAATCTCTCTCCATCACTTCCCTAGTCC 960

QY 961 ACTCTTGTGGCTGGGAACTTCTTGAACCTTAACTCTCTGCGAGCCCTTCTAAGACCCA 1020
DB 961 ACTCTTGTGGCTGGGAACTTCTTGAACCTTAACTCTCTGCGAGCCCTTCTAAGACCCA 1020

QY 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAAATAAATAAATGAAGAGGGGCAAA 1080
```

```
DB 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAAATAAATAAATGAAGAGGGGCAAA 1080
QY 1081 AAAAAAAAAAAAA 1091
DB 1081 AAAAAAAAAAAAA 1091

RESULT 3
HSM08731 Homo sapiens mRNA; cDNA DKFZp686H1078 (from clone DKFZp686H1078).
LOCUS Homo sapiens mRNA; cDNA DKFZp686H1078 (from clone DKFZp686H1078).
DEFINITION Homo sapiens mRNA; cDNA DKFZp686H1078 (from clone DKFZp686H1078).
ACCESSION BX648580
VERSION BX648580.1 GI:34367742
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Bloembergen, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686H1078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..811
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686H1078"
/tissue_type="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polyA_signal 702..707
polyA_site 723
ORIGIN
Query Match 46.8%; Score 510.8; DB 9; Length 811;
Best Local Similarity 99.6%; Pred. No. 1.1e-105;
Matches 512; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 578 ACCATATCCCGATCTGCTCCAGTGCCTCAACTCTCCATCTGCTCCCATGCGACCTGCC 637
DB 226 ACCATATCCCGATCTGCTCCAGTGCCTCAACTCTCCATCTGCTCCCATGCGACCTGCC 285

QY 638 ATGGTGTGATCCCGGAGAAATCAGAGCAACATGTTGTGCGAGCGCGCTCCCGGGGC 697
DB 286 ATGGTGTGATCCCGGAGAAATCAGAGCAACATGTTGTGCGAGCGCGCTCCCGGGGC 345

QY 698 AGGATGCTCTCCAGAGGTGATCTTGGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAAGTTC 757
DB 346 AGGATGCTCTCCAGAGGTGATCTTGGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAAGTTC 405

QY 758 TGGTGTCTGGGGTCTGTGGGGCCCCCTGTGGAACAAGATGGCAATCTCTGAGTCTACACT 817
DB 406 TGGTGTCTGGGGTCTGTGGGGCCCCCTGTGGAACAAGATGGCAATCTCTGAGTCTACACT 465

QY 818 ATATTGCAAGTATGTGGATCCGATGATCATGAGGACAACTACCTGTTTCTCT 877
DB 466 ATATTGCAAGTATGTGGATCCGATGATCATGAGGACAACTACCTGTTTCTCT 525
```

QY 878 CCACCTCCACCCCAACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCT 937
Db 526 CCACCTCCACCCCAACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCT 585
QY 938 CTTCCATCACTTCCCTAGCTAGCTCTTGTGGCTGGGAACCTTCTTGAACCTTTAACT 997
Db 586 CTTCCATCACTTCCCTAGCTAGCTCTTGTGGCTGGGAACCTTCTTGAACCTTTAACT 645
QY 998 CTTCCAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAGTGTGCAATAGTCTGGGAATA 1057
Db 646 CTTCCAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAGTGTGCAATAGTCTGGGAATA 705
QY 1058 AATATAATGAGAGGGGCAAAAAA 1091
Db 706 AATATAATGAGAGGGGCAAAAAA 739
RESULT 4
AR337568
LOCUS AR337568 840 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 5 from patent US 6566498.
ACCESSION AR337568
VERSION AR337568.1 GI:33723963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 840)
AUTHORS Ni, J. and Ruben, S. M.
TITLE Human serine protease and serpin polypeptides
JOURNAL Patent: US 6566498-A 5 20-MAY-2003;
FEATURES
source
1..840
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 46.4%; Score 506.4; DB 6; Length 840;
Best Local Similarity 97.9%; Pred. No. 1.1e-104;
Matches 513; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 568 CACCACGGAACCCATTCCCGGATCTGCTCAGTGCCTCAACCTCTCCATCGTCTCCCAT 627
Db 304 CAGCGGAGACCCATTCCCGGATCTGCTCAGTGCCTCAACCTCTCCATCGTCTCCCAT 363
QY 628 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 687
Db 364 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 423
QY 688 GTCCCGGGCAGAGTGCCTGCCAGGGTATCTCGGGGCCCCCTGGTGTGGGGGAGTC 747
Db 424 GTCCCGGGCAGAGTGCCTGCCAGGGTATCTCGGGGCCCCCTGGTGTGGGGGAGTC 483
QY 748 CTTCAAGTCTGGTGTCTCGGGGGTCTGTGGGGCCCCCTGGCAAGATGGCATCCCTGGA 807
Db 484 CTTCAAGTCTGGTGTCTCGGGGGTCTGTGGGGCCCCCTGGCAAGATGGCATCCCTGGA 543
QY 808 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGAACACTGA 867
Db 544 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGAACACTGA 603
QY 868 CTTGTTTCTCCACTCCACCCCAACCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGC 927
Db 604 CTTGTTTCTCCACTCCACCCCAACCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGC 663
QY 928 ACCAATATCTCTCCATCACTTCCCTAGCTCCATCTTGTGGCTGGGAATCTTTGG 987
Db 664 ACCAATATCTCTCCATCACTTCCCTAGCTCCATCTTGTGGCTGGGAATCTTTGG 723
QY 988 AACTTTACTCTGCGAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAGTGTGCAATA 1047
Db 724 AACTTTACTCTGCGAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAGTGTGCAATA 783

QY 1048 GTCTGGAATAAATATAAATGAAGGAGGGGCAAAAAA 1091
Db 784 GTCTGGAATAAATATAAATGAAGGAGGGGCAAAAAA 827
RESULT 5
BD130047
LOCUS BD130047 840 bp DNA linear PAT 18-SEP-2002
DEFINITION Human serine protease and serpin polypeptide.
ACCESSION BD130047
VERSION BD130047.1 GI:23224992
KEYWORDS JP 2002502600-A/3
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 840)
AUTHORS Ruben, S. M. and Ni, J.
TITLE Human serine protease and serpin polypeptide
JOURNAL Patent: JP 2002502600-A 3 29-JAN-2002;
COMMENT HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002502600-A/3
PD 29-JAN-2002
PF 04-FEB-1999 JP 2000530597
PR 06-FEB-1998 US 60/073961
PI STEVEN M RUBEN, JIAN NI
PC C12N9/64, A61K38/48, A61P7/02, A61P7/04, A61P7/06, A61P29/00, A61P31/ PC
00, A61P35/00, A61P37/06, C07K14/81, C07K16/40, C12N1/15, C12N1/19, PC
C12N1/21
PC C12N5/10, C12N15/09, C12Q1/37, A61K37/547, C12N5/00, C12N15/00 CC
Human serine protease and serpin polypeptide
FH Key Location/Qualifiers
FT CDS (115)..(603).
FEATURES
source
1..840
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 46.4%; Score 506.4; DB 6; Length 840;
Best Local Similarity 97.9%; Pred. No. 1.1e-104;
Matches 513; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 568 CACCACGGAACCCATTCCCGGATCTGCTCAGTGCCTCAACCTCTCCATCGTCTCCCAT 627
Db 304 CAGCGGAGACCCATTCCCGGATCTGCTCAGTGCCTCAACCTCTCCATCGTCTCCCAT 363
QY 628 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 687
Db 364 GCCACCTGCCATGGTGTATCCCGGAGAAATCAGAGCAACATGGTGTGCGAGGCGGC 423
QY 688 GTCCCGGGCAGAGTGCCTGCCAGGGTATCTCGGGGCCCCCTGGTGTGGGGGAGTC 747
Db 424 GTCCCGGGCAGAGTGCCTGCCAGGGTATCTCGGGGCCCCCTGGTGTGGGGGAGTC 483
QY 748 CTTCAAGTCTGGTGTCTCGGGGGTCTGTGGGGCCCCCTGGCAAGATGGCATCCCTGGA 807
Db 484 CTTCAAGTCTGGTGTCTCGGGGGTCTGTGGGGCCCCCTGGCAAGATGGCATCCCTGGA 543
QY 808 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGAACACTGA 867
Db 544 GTCTACACTATATTGCAAGTATGTGACTGGATCCGGATGATCATGAGAACACTGA 603
QY 868 CTTGTTTCTCCACTCCACCCCAACCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGC 927
Db 604 CTTGTTTCTCCACTCCACCCCAACCTTAACCTTGGGTACCCCTCTGGCCCTCAGAGC 663
QY 928 ACCAATATCTCTCCATCACTTCCCTAGCTCCATCTTGTGGCTGGGAATCTTTGG 987
Db 664 ACCAATATCTCTCCATCACTTCCCTAGCTCCATCTTGTGGCTGGGAATCTTTGG 723
QY 988 AACTTTACTCTGCGAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAGTGTGCAATA 1047
Db 724 AACTTTACTCTGCGAGCCCTTCTAAGACCCACAGCGGGGTGAGAGAGTGTGCAATA 783

Db 664 ACCAATATCTCTCCATCACTTCCCTAGCTCCACTCTTGTGGCTGGGAATCTTGG 723

QY 988 AACTTTAACTCTGCGCAGCCCTTCTAAGACCCACGAGCGGGTGAGAGAAGTGTGCAATA 1047

Db 724 AACTTTAACTCTGCGCAGCCCTTCTAAGACCCACGAGCGGGTGAGAGAAGTGTGCAATA 783

QY 1048 GTCTGGAATAAATATATAATGAAGGAGGGGCAAAAAAAAAAAAAA 1091

Db 784 GTCTGGAATAAATATATAATGAAGGAGGGGCAAAAAAAAAAAAAA 827

RESULT 6

LOCUS AR263961 619 bp DNA linear PAT 29-JAN-2003

DEFINITION Sequence 139 from patent US 6331427.

ACCESSION AR263961

VERSION AR263961.1 GI:28075965

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 619)

AUTHORS Robison,K.E.

TITLE Protease homologs

JOURNAL Patent: US 6331427-A 139 18-DEC-2001;

FEATURES

Location/Qualifiers

1..619

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 45.2%; Score 493.2; DB 6; Length 619;

Best Local Similarity 97.7%; Pred. No. 1.1e-101;

Matches 543; Conservative 0; Mismatches 8; Indels 5; Gaps 4;

QY 98 CAGACTTTGGAAGTGAACCCACCATGGGCTCAGCATCTTTTGTCTCTGTGTCTTTG 157

Db 67 CAAGACTTTGGAAGTGAACCCACCATGGGCTCAGCATCTTTTGTCTCTGTGTCTTTG 126

QY 158 GACTCAGCCAGGCGGACCCACCGAGATTTTCAATGGCACTGAGTGTGGCGTAACTCAC 217

Db 127 GACTCAGCCAGGCGGACCCACCGAGATTTTCAATGGCACTGAGTGTGGCGTAACTCAC 186

QY 218 AGCCGTGGCAGGTGGGGCTGTTTGAAGGACCCAGCCCTGCGCTGCGGGGTGTCCTTATTG 277

Db 187 AGCCGTGGCAGGTGGGGCTGTTTGAAGGACCCAGCCCTGCGCTGCGGGGTGTCCTTATTG 246

QY 278 ACCAAGTGGGTCTCAGAGGGCTCAGAGGGCTCAGAGGGCTCAGAGGGCTCAGAGGGCT 337

Db 247 ACCAAGTGGGTCTCAGAGGGCTCAGAGGGCTCAGAGGGCTCAGAGGGCTCAGAGGGCT 306

QY 338 GGGACACAGCTCAGCAGCTCAGTGTGACCGAGCAGATCCGACAGCGGCTTCTCTG 397

Db 307 GGGACACAGCTCAGCAGCTCAGTGTGACCGAGCAGATCCGACAGCGGCTTCTCTG 366

QY 398 TGACCCATCCCGCTACCTGGAGGCTTCGACGAGCCACGAGCAGCAGCTCCGCTGCTGC 457

Db 367 TGACCCATCCCGCTACCTGGAGGCTTCGACGAGCCACGAGCAGCAGCTCCGCTGCTGC 426

QY 458 GACTGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 517

Db 427 GACTGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 486

QY 518 GTGCAACCGCTGCGCAGGTGCGAGTCTCAGGCTGGGGGATCACCAACCCACCGCA 577

Db 487 GTGCAACCGCTGCGCAGGTGCGAGTCTCAGGCTGGGGGATCACCAACCCACCGCA 546

QY 578 ACCAATCCCGATGTCTCCAGTGTCTCAACCTCTCCATCGTCTCCATGCGACCTGCGC 637

Db 547 ACCAATTC -GGATCTGTCCAGTGTCTCAACCTCT -CATGCTCT -CCAATGCCACTG -C 601

QY 638 ATGGTGTGTATCCCG 653

Db 602 ATGGTGTGTATCCGG 617

RESULT 7

LOCUS AF164623 9120 bp DNA linear PRI 26-JUN-2000

DEFINITION Homo sapiens trypsin-like serine protease (TLSP) gene, complete cds.

ACCESSION AF164623

VERSION AF164623.1 GI:5713130

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 9120)

AUTHORS Yousef,G.M., Scorilas,A. and Diamandis,E.P.

TITLE Genomic organization, mapping, tissue expression, and hormonal regulation of trypsin-like serine protease (TLSP PRSS20), a new member of the human kallikrein gene family

JOURNAL Genomics 63 (1), 88-96 (2000)

MEDLINE 20130117

PUBMED 10662548

REFERENCE 2 (bases 1 to 9120)

AUTHORS Yousef,G.M., Scorilas,A. and Diamandis,E.P.

TITLE Direct Submission

JOURNAL Submitted (01-JUL-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1x5, Canada

FEATURES

Location/Qualifiers

1..9120

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="19"

/map="19q13.3-q13.4"

2313..7622

/gene="TLSP"

Join(2313..2398,4189..4263,5061..5217,5545..5810,6627..6763,7158..7622)

/gene="TLSP"

/product="trypsin-like serine protease"

Join(4224..4263,5061..5217,5545..5810,6627..6763,7158..7310)

/gene="TLSP"

/note="kallikrein-like serine protease"

/codon_start=1

/product="trypsin-like serine protease"

/protein_id="AAD47815.1"

/db_xref="GI:5713131"

/translation="MRILQLILLALATGLVGGTFLIKGFBCPKHSPQWQALFKTRLLCGATLAPRMLLTAAHCLPRYIVHGLQNNLQKEGCEQIRATFSPFHPGFNNSLFNKDRNIMLVKMASPVSTIWAVRPLTSSRCVTAGTSLTSGWGSTSPQLRLPHTLRCANITTIHQKCNAYPGNITDTMVCASQVEGGKSCQSDSGGPLVCNOSLQIIISWGQDPCATIRKPGVYTKVKYVDWIQTMMKN"

ORIGIN

Query Match 34.2%; Score 373.2; DB 9; Length 9120;

Best Local Similarity 99.2%; Pred. No. 2e-74;

Matches 375; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 702 TGCCCTCCAGGGTGTATCTGGGGGCCCTGTGTGTGGGGAGTCCCTCAAGGCTGTGT 761

Db 388 TGCCCTCCAGGGTGTATCTGGGGGCCCTGTGTGTGGGGAGTCCCTCAAGGCTGTGT 447

QY 762 GTCTGGGGGTGTGTGGGGGCCCTGTGGAACAAGATCCCTGGAGTCTACACCTATAT 821

Db 448 GTCTGGGGGTGTGTGGGGGCCCTGTGGAACAAGATCCCTGGAGTCTACACCTATAT 507

QY 822 TTGCAAGTATGTGGAAGTGTGATCCGATGATCATGAGGAACAAGTCTGTTCTCCAC 881

Db 508 TTGCAAGTATGTGGAAGTGTGATCCGATGATCATGAGGAACAAGTCTGTTCTCCAC 567

```

QY      882  CTCACCCCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTC 941
Db      568  CTCACCCCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTC 627
QY      942  CATCACTTCCCTAGCTCCACTCTTGTGGCCCTGGGAACCTCTTGGAACCTTAACTCTCTG 1001
Db      628  CATCACTTCCCTAGCTCCACTCTTGTGGCCCTGGGAACCTCTTGGAACCTTAACTCTCTG 687
QY      1002 CCAGCCCTTAAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATA 1061
Db      688  CCAGCCCTTAAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATA 747
QY      1062 TAAATGAAGGAGGGGCAA 1079
Db      748  TAAATGAAGGAGGGGCCA 765

RESULT 8
LOCUS   AF135025
DEFINITION Homo sapiens kallikrein-like protein 5 gene, alternative splice products, complete cds.
ACCESSION AF135025
VERSION   AF135025
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS   Yousef,G.M., Luo,L. and Diamandis,E.P.
TITLE      Identification of novel human kallikrein-like genes on chromosome
           19q13.3-q13.4
JOURNAL   Anticancer Res. 79, 2843-2852 (1999)
REFERENCE
AUTHORS   Diamandis,E.P., Yousef,G.M., Luo,L.Y., Magklara,A. and Obiezu,C.V.
TITLE      The new human kallikrein gene family: implications in
           carcinogenesis
JOURNAL   Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
MEDLINE   21121728
PUBMED    10675891
REFERENCE
AUTHORS   Yousef,G.M., Magklara,A. and Diamandis,E.P.
TITLE      KLK-L5 is a novel serine protease and a new member of the human
           kallikrein gene family-differential expression in breast cancer
           Unpublished
JOURNAL   (bases 1 to 11820)
REFERENCE
AUTHORS   Yousef,G.M., Magklara,A., Scorilas,A. and Diamandis,E.P.
TITLE      Cloning of new alternatively spliced forms of the kallikrein-like
           gene 5 (KLK-L5)
JOURNAL   Unpublished
REFERENCE
AUTHORS   Yousef,G.M., Magklara,A. and Diamandis,E.P.
TITLE      Direct Submission
JOURNAL   Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
           Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
           Canada
REFERENCE
AUTHORS   Yousef,G.M., Magklara,A. and Diamandis,E.P.
TITLE      Direct Submission
JOURNAL   Submitted (01-NOV-1999) Pathology and Laboratory Medicine, Mount
           Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
           Canada
REMARK    Sequence update by submitter
COMMENT   On Nov 1, 1999 this sequence version replaced gi:4589276.
FEATURES
         location/Qualifiers
           1..11820
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="19"
             /map="19q13.3-q13.4"

```

```

mRNA
join(3743..3840,3995..4050,4496..4655,6500..6759,
7714..7847,9178..9543)
/product="kallikrein-like protein 5"
join(3743..3840,3995..4050,4496..4655,6500..6759,
7714..7847,9178..9293,9423..9543)
/product="kallikrein-like protein 5-related protein 1"
join(3743..3840,3995..4050,4496..4655,7714..7847,
9178..9543)
/product="kallikrein-like protein 5-related protein 2"
join(4014..4050,4496..4655,6500..6759,7714..7847,
9178..9293,9423..9480)
/note="KLK-L5-related protein 1; kallikrein-like serine
protease; alternatively spliced"
/codon_start=1
/product="kallikrein-like protein 5-related protein 1"
/protein_id="AAF06065.1"
/db_xref="GI:6249632"
/translation="MGLSIFLLLCVLGLSOAATPKIENGTECGRNSOPWQVGLPEGNS
LRGGVLIDHRWVLTAAHCSGRYVPLGHSLSQLDWTQIRHSGFSVTHPGVLGAS
TSHEHDRLRLRLPVRTSVQPLPNDCAATAGTSCVSGWGTINHPNPFDDLIQ
CLNLSIVSHATCHGVYECRITSNMVCAGVPGQDACQDGGPLVCGVLQGLVSWGS
VPCGQDGPVYTYICNLTGLGTSWNFNSCQPF"
join(4014..4050,4496..4655,6500..6759,7714..7847,
9178..9333)
/note="KLK-L5"
/codon_start=1
/product="kallikrein-like protein 5"
/protein_id="AAD26426.2"
/db_xref="GI:6166249"
/translation="MGLSIFLLLCVLGLSOAATPKIENGTECGRNSOPWQVGLPEGNS
LRGGVLIDHRWVLTAAHCSGRYVPLGHSLSQLDWTQIRHSGFSVTHPGVLGAS
TSHEHDRLRLRLPVRTSVQPLPNDCAATAGTSCVSGWGTINHPNPFDDLIQ
CLNLSIVSHATCHGVYECRITSNMVCAGVPGQDACQDGGPLVCGVLQGLVSWGS
VPCGQDGPVYTYICNLTGLGTSWNFNMRNN"
join(4014..4050,4496..4655,7714..7847,9178..9182)
/note="KLK-L5-related protein 2; kallikrein-like serine
protease; alternatively spliced"
/codon_start=1
/product="kallikrein-like protein 5-related protein 2"
/protein_id="AAF06066.1"
/db_xref="GI:6249633"
/translation="MGLSIFLLLCVLGLSOAATPKIENGTECGRNSOPWQVGLPEGNS
LRGGVLIDHRWVLTAAHCSGRIPGPSAPVPLRLPCHLPWCVSRNHEQHVCR
RPGAGCPLG"
polyA_signal 9520..9525
ORIGIN
Query Match 34.2%; Score 373.2; DB 9; Length 11820;
Best Local Similarity 99.2%; Pred. No. 2e-74;
Matches 375; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 702 TGCCTGCCAGGGTGATCTGGGGCCCCCTGGTGTGTGGGGAGTCCTTCAAGTCTGGT 761
Db 9168 TCCCTTCAGGGTGATCTGGGGCCCCCTGGTGTGTGGGGAGTCCTTCAAGTCTGGT 9227
QY 762 GTCTCGGGGTCTGTGGGGCCCTGTGCACAGATGGCATCCCTCGAGTCTACACCTATAT 821
Db 9228 GTCTCGGGGTCTGTGGGGCCCTGTGCACAGATGGCATCCCTCGAGTCTACACCTATAT 9287
QY 822 TTGCAAGTATGTGAGTCGGATCCGGATGATCATCAGGAACAACACTGCTGTTCTCCAC 881
Db 9288 TTGCAAGTATGTGAGTCGGATCCGGATGATCATCAGGAACAACACTGCTGTTCTCCAC 9347
QY 882 CTCCACCCCAACCCCTTAACCTGGGTACCCCTCTGGGCCCTCAGAGCAACATATCTCTC 941
Db 9348 CTCCACCCCAACCCCTTAACCTGGGTACCCCTCTGGGCCCTCAGAGCAACATATCTCTC 9407
QY 942 CATCACTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTCTTGGAACTTTAACTCTCTG 1001
Db 9408 CATCACTTCCCTAGCTCCACTCTTGTGGCTGGGAACCTCTTGGAACTTTAACTCTCTG 9467
QY 1002 CCAGCCCTTAAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATA 1061

```


	Query Match	34.2%;	Score 373.2;	DB 9;	Length 132323;
	Best Local Similarity	99.2%;	Pred. No. 1.8e-74;		
	Matches 375;	Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
QY	702	TGCTGCCAAGGTGATTCGGGGGGCCCCCGTGTGTGTGGGGAGTCCTCAAGGTCCTGGT	761		
dDb	29548	TCCCTTCACAAGGTGATTCGGGGGGCCCCCGTGTGTGTGGGGAGTCCTCAAGGTCCTGGT	29489		
QY	762	GPCCTGGGGGGTCTGTGGGGCCCCGTGTGGACAAGATGGCATCCCTCGAGTGCTACACTATAT	821		

IEPNFSPDDLLQCVDLKILPNDECKAHVQKVDTFMLCVGHLEGGKDTCVGDSGGPL
MCDGLVQGVTSWGYVPCGTPNKPSPVAVRLSVXWIEDTIAENS"
complement(join(<13552. .13704,14377. .14530,14678. .14917,
15416. .15569,19204. .>19246))
/product="ACO protease"
complement(join(13552. .13704,14377. .14530,14678. .14917,
15416. .15569,19204. .19246))
/note="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MWLVLTSLFLASTAAQDGDKLLEGDECAPIHSQWQVVALYERGR
FNCASLISPHWLVSAACQCFMRVRLGEHLNRKRDGPBQLRTTSVIPHRYEARS
HRNDIMLRVQPARLNPQVPAVLPTCPHPGEACVVGWGLVPLSPVSLPDTLHC
ANISITSQCDKSYGRLLTNTMVCAGAEGRGAECEGDSGGPLVCGGILQGIIVSGD
VPCDNTTGPVGVTVKCHYLEWIRETKEN"
/gene="KLK3"
/join(<42595. .42640,43880. .44039,45669. .45955,46099. .46235,
47614. .>47769)
/gene="KLK3"
/product="prostate specific antigen"
join(42595. .42640,43880. .44039,45669. .45955,46099. .46235,
47614. .47769)
/gene="KLK3"
/note="serine protease"
/codon_start=1
/product="prostate specific antigen"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translation="MWVPPVFTLSVTWIGAAPILSRIKRVGVGWECEKHSQWQVILVAS
RGRVCGVIVHPQWVLTAAHCIRKNSVILGRHSLFHPEDTGQVQVSHSFPHPLVD
MSLLKRLRPEDSDHMLRLSEPAELTDVAKVMDLPTQEPALGTTCVASGWSI
EPPEFLTPKQCVDLHVISNDVCAQHPQKVTFMLCAGRTGSKTSCSDSGGLV
CNGVLQGITSGSPCALPERPSLYTKVHVHKKIKTIVANP"
/gene="KLK2"
/join(<61139. .61184,62391. .62550,64142. .64428,64542. .64678,
66074. .>66229)
/gene="KLK2"
/product="glandular kallikrein"
join(61139. .61184,62391. .62550,64142. .64428,64542. .64678,
66074. .66229)
/gene="KLK2"
/note="serine protease"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MWDLVLIALSVGCTGAVPLIQSRIVGVGWECEKHSQWQVAVYS
HGWAHCGVIVHPQWVLTAAHCIRKNSQWILGRNLFEPEDTGQVQVSHSFPHPLYN
MSLLKHSRLPDESDHMLRLSEPAKITDVVKVLGLPTQEPALGTTCVASGWSI
EPPEFLRPSLQCVSLHSLNDMCARAYSEKVTFMLCAGITGSKTSCSDSGGLV
CNGVLQGITSGPPCALPEKPAVYTKVHVHKKIKTIVANP"
complement(join(<94602. .94754,96027. .96163,96247. .96497,
96919. .97081,98345. .>98405))
/product="serine protease"
complement(join(94602. .94754,96027. .96163,96247. .96497,
96919. .97081,98345. .98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPWGWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPW
QAALWNEFLPCSGVLPWQVLSAACHFQNSYITIGLHSLLEADQPSQWFEASLS
VRPEXNPLIANDMLIKLDESSEDYTESISIAQCTPAGNSCLVSGWGLLANGER
MPLVLCQNVSVSEVCSKLYDLRPSMFCAGGHDQKSDGSDSGGLICNGYILO
GLVFGKAPCCQGVGPGVYTNLCKFTWIEKTQAS"
complement(<131301. .>140303)
/gene="KLK5"
/note="synonym: SCTE"
complement(join(<131301. .131456,136310. .136443,

136529. .136785,137525. .137690,140255. .>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement(join(131301. .131456,136310. .136443,
136529. .136785,137525. .137690,140255. .140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWVLCALITALLGVTDARSDSSRIINGSDDMHTQPMQAAAL
LLRNLQYCGAVLPQWMLTAAHCKKRVFRLGHVLSVYSGQQMFGVKSIPIH
PGYHPGNSDMLIKLNRIRPTKDRPINVSHGCSAGTKLVSNGTTSKSPVHF
PKVLCQNLISVLSCRGEDAVPRDIDTMFCAGDKAGRDSQQGDSGGPVVCSGLQGL
VSGWDYPCARNRGGVYTNLCKFTWIEKTQAS"
complement(join(<146834. .146986,149628. .149764,
151186. .151433,155052. .155208,155948. .>155987))
/product="protease M"
complement(join(146834. .146986,149628. .149764,
151186. .151433,155052. .155208,155948. .155987))
/note="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKKLMVLSLIAAAWAEQNLVHGGPCDKTSHPYOALYTSGH
LLCGVLILHPLVLTAAHCKKPNLQVFLGKHLNLRQSSSQSSVRAVHPDADAAS
HDQDMLLRPLARPAKLSLQPLELRDCSNTSCHILGWKTADGFPPTIOCAIY
HLVSRCEHAYPGQITONMLCAGDEKYKQDSQCGSGPLVCGDHLRGLVSWGNIPC
GSKEPGVYTNVCRYTNWIKTIQAK"
complement(<165420. .>170283)
/gene="KLK7"
/note="synonym: SCCE"
complement(join(<165420. .165575,167672. .167808,
168124. .168371,169651. .169798,170211. .>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement(join(165420. .165575,167672. .167808,
168124. .168371,169651. .169798,170211. .170283))
/gene="KLK7"
/note="serine protease"
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MABSLPLQLILLSLALETAGEAQGDKIIDGAPCARGSHPWQ
VALLSQNLHCGVLVNERVLTAAHCKMNEYVHLGSDITGDRRAQRIKASKSRHP
GYSTQTHVNDMLVUNSOARLSMVKKRLPSRCEPPTCTVSGWGTTSPDVTFF
SDLMCDVKLISPDCTKYKDLLENMLCAGIPDSKKNACNGSDGGLVCRGTLOGL
VSWGTFPGQNPDPGVYTVCKFTKINDTMKKHR"
complement(join(<183943. .184098,185635. .185768,
187865. .188127,188293. .188452,188967. .>189036))
/product="neuropsin"
complement(join(183943. .184098,185635. .185768,
187865. .188127,188293. .188452,188967. .189036))
/note="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRPRAAKTMWFLLLGGAWAGHRAQEDKVLGGHECOPHSQ
PWQAALFQQLCGVLVGNVLTAAHCKKPKYTVRLGHSIQNDKQDGEQELPVVQ
SIPPCVNSDSDVDHNDMLLQRLQASGSKVKPISLADHTCQPKCTVSGWGT
TSRENFEDPLNCAEVKIPQKCKEDAYPQITDGMVCAGSSKAGATCCQDSDGGPLVC
DGAQGITSGWSPCGRSDKPGVYTNICRILDMIKIIGSKG"
complement(join(<190980. .191129,191573. .191709,
194324. .194589,197048. .197204,197370. .>197412))
/product="kallikrein-like 3"
complement(join(190980. .191129,191573. .191709,
194324. .194589,197048. .197204,197370. .197412))
/note="serine protease"
/codon_start=1
/product="kallikrein-like 3"


```

/notes="clone overlaps with GenBank Accession Number
AC130782 clone CH251-355A20 (center project name dhz)"
misc_feature 1..47270
/notes="assembly_fragment
clone_end:17
vector_side:left"
misc_feature 47371..58115
/notes="assembly_fragment"
misc_feature 58216..61050
/notes="assembly_fragment"
misc_feature 61151..65346
/notes="assembly_fragment"
misc_feature 65447..71536
/notes="assembly_fragment"
misc_feature 71637..83139
/notes="assembly_fragment"
misc_feature 83240..85133
/notes="assembly_fragment"
misc_feature 85234..88837
/notes="assembly_fragment"
misc_feature 88938..129108
/notes="assembly_fragment"
misc_feature 113257..174724
/notes="clone overlaps with GenBank Accession Number
AC133392 clone CH251-126O24 (center project name dxa)"
misc_feature 129209..174724
/notes="assembly_fragment
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 33.8%; Score 368.4; DB 2; Length 174724;
Best Local Similarity 98.4%; Pred. No. 2.2e-73;
Matches 372; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 702 TGCTGCGGAGGTGATTCTGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGTCTGGT 761
Db 38231 TCCTTCCAGGTGATTCTGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGTCTGGT 38172
QY 762 GTCCTGGGGCTCTGGGGCCCTGTGGACAAGATGTCATCCCTGGAGTCTACACCTATAT 821
Db 38171 GTCCTGGGGCTCTGGGGCCCTGTGGACAAGATGTCATCCCTGGAGTCTACACCTATAT 38112
QY 822 TTGCAAGTATGTGAGTATGCGATGATCATGAGAACTGACCTGTTTCCCTCCAC 881
Db 38111 TTGCAAGTATGTGAGTATGCGATGATCATGAGAACTGACCTGTTTCCCTCCAC 38052
QY 882 CTCACCCGCCACCCCTTAATCTGGGTACCCCTCTGGCCCTCAGACCAATATCTCCTC 941
Db 38051 CTCACCCGCCACCCCTTAATCTGGGTACCCCTCTGGCCCTCAGACCAATATCTCCTC 37992
QY 942 CATCACTTCCCTTAGCTCCACTCTGTGGCCCTGGGAATCTTGGAACTTTAACTCCTG 1001
Db 37991 CATCACTTCCCTTAGCTCCACTCTGTGGCCCTGGGAATCTTGGAACTTTAACTCCTG 37932
QY 1002 CCAGCCCTTCTAAGACCCAGCGGGGTGAGAAAGTGTGCAATAGTCTGGAATAAATA 1061
Db 37931 CCAGCCCTTCTAAGACCCAGCGGGGTGAGAAAGTGTGCAATAGTCTGGAATAAATA 37872
QY 1062 TAAATGAAGGGGGCAA 1079
Db 37871 TAAATGAAGGGGGCAA 37854

RESULT 12
AC130782/c
LOCUS
DEFINITION Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
ordered pieces.
AC130782
AC130782.2 GI:25167101
VERSION
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Pan troglodytes (Chimpanzee)

```

ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE
AUTHORS

1 (bases 1 to 200792)

Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-O., Legaspi,R., Maskeri,B., McDowell,J., Margulies,E.H., Masiello,C., Pearson,R., Portnoy,M.B., Prasad,A., Paguirigan,C., Pearson,R., Schandler,K., Schueler,M.G., Sison,C., Reddix-Dugue,N., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 200792)

Green,E.D.

Direct Submission

Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 200792)

Green,E.D.

Direct Submission

Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717

Grovmont Circle, Gaithersburg, MD 20877, USA

On Nov 22, 2002 this sequence version replaced gi:22218452.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc.zoo@nhgri.nih.gov

----- Project Information

Center project name: dhz

Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 198656 bases at least Q40

Consensus quality: 197883 bases at least Q30

Consensus quality: 198879 bases at least Q20

Insert size: 215000; agarose-fp

Quality coverage: 9.03x in Q20 bases; agarose-fp

Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 28306: contig of 28306 bp in length

* 28307 28406: gap of unknown length

* 28407 37856: contig of 9450 bp in length

* 37857 37956: gap of unknown length

```

* 37957 73522: contig of 35566 bp in length
* 73523 73622: gap of unknown length
* 83567 83567: contig of 9945 bp in length
* 83668 83667: gap of unknown length
* 88818 88817: contig of 5150 bp in length
* 88918 88917: gap of unknown length
* 88918 125611: contig of 36694 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174798: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193444: contig of 6862 bp in length
* 193445 193444: gap of unknown length
* 193445 199363: contig of 5919 bp in length
* 199364 199463: gap of unknown length
* 199464 200792: contig of 1329 bp in length.

```

FEATURES

```

source
1..200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"

```

misc_feature

```

1..28306
/note="assembly_fragment
clone_end:SP6
vector_side:left"

```

misc_feature

```

28407..37856
/note="assembly_fragment"

```

misc_feature

```

37957..73522
/note="assembly_fragment"

```

misc_feature

```

73623..83567
/note="assembly_fragment"

```

misc_feature

```

83668..88817
/note="assembly_fragment"

```

misc_feature

```

88918..125611
/note="assembly_fragment"

```

misc_feature

```

125712..159879
/note="assembly_fragment"

```

misc_feature

```

159980..174698
/note="assembly_fragment"

```

misc_feature

```

174799..186382
/note="assembly_fragment"

```

misc_feature

```

186483..193444
/note="assembly_fragment"

```

misc_feature

```

193445..199363
/note="assembly_fragment"

```

misc_feature

```

199464..200792
/note="assembly_fragment
clone_end:17
vector_side:right"

```

ORIGIN

```

Query Match      33.8%; Score 368.4; DB 2; Length 200792;
Best Local Similarity 98.4%; Pred. No. 2.e-73;
Matches 372; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 702 TGCCTGCCAGGTGATCTTGGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGGTCGTGT 761
DB 148663 TCCCTTCCAGGGTGATCTTGGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGGTCGTGT 148604
QY 762 GTCCTGGGGGCTCTGGGGGCCCCCTGGCAAGATGGCATCCCTGGAGTCTACACCTATAT 821
DB 148603 GTCCTGGGGGCTCTGGGGGCCCCCTGGCAAGATGGCATCCCTGGAGTCTACACCTATAT 148544
QY 822 TTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACCTGATCTTCTCCAC 881
DB 148543 TTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAACCTGATCTTCTCCAC 148484
QY 882 CTCACCCCCCAACCCCTTAACCTTGGGGTACCCCTCTGGCCCTCAGACCAATATCTCCTC 941

```

```

Db 148483 CTCACCCCCCACCCTTAGCTTGGGTACCACTCTGGCCCCCAGAGACCAATATCTCCTC 148424
QY 942 CATCACTTCCCTTAGCTTCCACTCTTGTGGCCTGGGAACCTTCTTGGAACTTTAACTCCTG 1001
DB 148423 CATCACTTCCCTTAGCTTCCACTCTTGTGGCCTGGGAACCTTCTTGGAACTTTAACTCCTG 148364
QY 1002 CGAGCCCTTCTTAAGACCCAGGAGGGGGTGGAGAGTGTGCAATAGTCTGGAAATAATA 1061
DB 148363 CGAGCCCTTCTTAAGACCCAGGAGGGGGTGGAGAGTGTGCAATAGTCTGGAAATAATA 148304
QY 1062 TAAATGAAGGAGGGGCAA 1079
DB 148303 TAAATGAAGGAGGGGCAA 148286

RESULT 13
AC073185/c
LOCUS AC073185 142334 bp DNA linear HTG 17-AUG-2000
DEFINITION Homo sapiens chromosome 19 clone RP11-10111, WORKING DRAFT
SEQUENCE, 32 unordered pieces.
ACCESSION AC073185
VERSION AC073185.4 GI:9838034
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142334)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 142334)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Aug 17, 2000 this sequence version replaced gi:9653152.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH001011
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: bases at least Q40
Consensus quality: bases at least Q30
Consensus quality: bases at least Q20
Insert size: 135000; agarose-fp
Insert size: 139234; sum-of-contigs
Quality coverage: 3.86 in Q20 bases; agarose-fp
Quality coverage: 3.87 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1452: contig of 1452 bp in length
* 1453 1552: gap of unknown length
* 1553 2737: contig of 1185 bp in length
* 2738 2837: gap of unknown length
* 2838 4665: contig of 1828 bp in length

```

```

* 4666 4765: gap of unknown length
* 4766 contig of 1289 bp in length
* 6054 6154: gap of unknown length
* 6055 contig of 1637 bp in length
* 6155 7791: gap of unknown length
* 7792 10026: contig of 2135 bp in length
* 7892 10127: gap of unknown length
* 10027 11895: contig of 1769 bp in length
* 11896 11996: gap of unknown length
* 11997 15171: gap of unknown length
* 15172 18476: contig of 3075 bp in length
* 18477 18476: contig of 3306 bp in length
* 18477 18576: gap of unknown length
* 18577 21537: contig of 2961 bp in length
* 21538 21637: gap of unknown length
* 21638 24626: contig of 2988 bp in length
* 24627 27583: gap of unknown length
* 27584 27683: contig of 2858 bp in length
* 27684 30892: gap of unknown length
* 30893 30992: contig of 3209 bp in length
* 30993 35210: gap of unknown length
* 35211 39361: gap of unknown length
* 39362 42356: contig of 2995 bp in length
* 42357 42456: gap of unknown length
* 42457 46421: contig of 3965 bp in length
* 46422 51157: gap of unknown length
* 51158 51257: gap of unknown length
* 51258 57416: contig of 6159 bp in length
* 57417 61710: contig of 4194 bp in length
* 61711 66249: gap of unknown length
* 66250 66350: contig of 4439 bp in length
* 66351 71980: gap of unknown length
* 71981 72080: contig of 5631 bp in length
* 72081 77454: gap of unknown length
* 77455 84318: gap of unknown length
* 84319 88649: contig of 6764 bp in length
* 88650 88749: gap of unknown length
* 88750 94419: contig of 4231 bp in length
* 94420 94519: gap of unknown length
* 94520 100165: contig of 5670 bp in length
* 100166 100266: gap of unknown length
* 100267 107334: contig of 5846 bp in length
* 107335 107435: gap of unknown length
* 107436 115093: contig of 7069 bp in length
* 115094 115193: gap of unknown length
* 115194 122023: contig of 7659 bp in length
* 122024 122124: gap of unknown length
* 122125 132649: contig of 6830 bp in length
* 132650 132750: gap of unknown length
* 132751 142334: contig of 10526 bp in length
* 142335 142334: contig of 9585 bp in length.

```

FEATURES

Location/Qualifiers

```

1..142334
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-10111"
1..1452
/note="assembly_name:Contig26"
1553..2737
/note="assembly_name:Contig27"
2838..4665
/note="assembly_name:Contig28"
4766..6054
/note="assembly_name:Contig29"

```

misc_feature

misc_feature

misc_feature

misc_feature

```

misc_feature 6155..7791
/note="assembly_name:Contig30"
misc_feature 7892..10026
/note="assembly_name:Contig31"
misc_feature 10127..11895
/note="assembly_name:Contig33"
misc_feature 11996..15070
/note="assembly_name:Contig34"
misc_feature 15171..18476
/note="assembly_name:Contig35"
misc_feature 18577..21537
/note="assembly_name:Contig36"
misc_feature 21638..24625
/note="assembly_name:Contig37"
misc_feature 24726..27583
/note="assembly_name:Contig38"
misc_feature 27684..30892
/note="assembly_name:Contig39"
misc_feature 30993..35209
/note="assembly_name:Contig40"
misc_feature 35310..39261
/note="assembly_name:Contig41"
misc_feature 39362..42356
/note="assembly_name:Contig42"
misc_feature 42457..46421
/note="assembly_name:Contig43"
misc_feature 46522..51157
/note="assembly_name:Contig44"
misc_feature 51258..57416
/note="assembly_name:Contig45"
misc_feature 57517..61710
/note="assembly_name:Contig46"
misc_feature 61811..66249
/note="assembly_name:Contig47"
misc_feature 66350..71980
/note="assembly_name:Contig48"
misc_feature 72081..77454
/note="assembly_name:Contig49"
misc_feature 77555..84318
/note="assembly_name:Contig50"
misc_feature 84419..88649
/note="assembly_name:Contig51"
misc_feature 88750..94419
/note="assembly_name:Contig52"
misc_feature 94520..100165
/note="assembly_name:Contig53"
misc_feature 100266..107334
/note="assembly_name:Contig54"
misc_feature 107435..115093
/note="assembly_name:Contig55"
misc_feature 115194..122023
/note="assembly_name:Contig56"
misc_feature 122124..132649
/note="assembly_name:Contig57"
clone_end:r7
vector_side:right"
132750..142334
/note="assembly_name:Contig58"

```

ORIGIN

```

Query Match 33.1%; Score 361.2; DB 2; Length 142334;
Best Local Similarity 98.9%; Pred. No. 9.6e-72;
Matches 374; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 702 TGCTGCGAGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGAGTCCTTCAAGTCTGGT 761
Db 10990 TCCCTTCAGGGTATTCTGGGGGCCCCCTGGTGTGTGGGGGAGTCCTTCAAGTCTGGT 10931

Qy 762 GTCTGCGGGGTCTGTGGGGGCCCTGTGGACAAGATGGGATCCCTGAGTCTACACCTATAT 821
Db 10930 GTCTGCGGGGTCTGTGGGGGCCCTGTGGACAAGATGGGATCCCTGAGTCTACACCTATAT 10871

Qy 822 TTGCAAGTATGTGGATCGGATCCGGATGATCATGAGGACAACATGACCTGTTTCTCCAC 881

```

Mon Jun 21 09:00:25 2004

Db	10870	TTGCAAGTATGTGGACCTGGATCCGATCATATGAGGAACAACCTGACCTGTTTCTCCAC	10811
Qy	882	CTCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTC	941
Db	10810	CTCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTC	10752
Qy	942	CATCACTTCCCTAGCTCCACTCTTGTGGCTGGAACTTTTGAATTTAACTCTCTG	1001
Db	10751	CATCACTTCCCTAGCTCCACTCTTGTGGCTGGAACTTTTGAATTTAACTCTCTG	10692
Qy	1002	CCAGCCCTTAAAGCCACGAGCGGGTGAGAGTGTGCAATAGTCTGCAATAAATA	1061
Db	10691	CCAGCCCTTAAAGCCACGAGCGGGTGAGAGTGTGCAATAGTCTGCAATAAATA	10632
Qy	1062	TAAATGAAGAGGGGCA 1079	
Db	10631	TAAATGAAGAGGGGCA 10614	
RESULT 14			
AX070938			
LOCUS	AX070938	405 bp	DNA
DEFINITION	Sequence 1410 from Patent WO0102568.		linear
ACCESSION	AX070938		PAT 25-JAN-2001
VERSION	AX070938.1	GI:12581211	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D., Lamson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.		
TITLE	Human genes and gene expression products		
JOURNAL	Patent: WO 0102568-A 1410 11-JAN-2001;		
FEATURES	CHIRON CORPORATION (US) ; HYSEQ, INC. (US)		
source	Location/Qualifiers		
	1..405		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	27.2%	Score 297;	DB 6; Length 405;
Best Local Similarity	99.7%	Pred. No. 5.2e-57;	
Matches	308; Conservative	0; Mismatches	0; Indels
			1; Gaps
Qy	10	CATCCCTTGGTGGACCTCAAGAGAGCAGAGAGGCGAGAGGTTGGGGGACACAGGAA	69
Db	9	CATCCCTTGGTGGACCTCAAGAGAGCAGAGAGGCGAGAGGTTGGGGGACACAGGAA	68
Qy	70	GGGTGACCTCTGAGATCCCTTTCCTCCAGAGCTTTTGAAGTGACCCCATGGGCT	128
Db	69	GGGTGACCTCTGAGATCCCTTTCCTCCAGAGCTTTTGAAGTGACCCCATGGGCT	128
Qy	129	CAGCATCTTTTGTCTGTGTCTTGTGGCTTGAGCCAGGAGCCACCGAGATTTT	188
Db	129	CAGCATCTTTTGTCTGTGTCTTGTGGCTTGAGCCAGGAGCCACCGAGATTTT	188
Qy	189	CAATGGCACTGAGTGTGGCGTAACTCAAGCCGTTGGAGTGGGGCTTTTGGGGCAC	248
Db	189	CAATGGCACTGAGTGTGGCGTAACTCAAGCCGTTGGAGTGGGGCTTTTGGGGCAC	248
Qy	249	CAGCCTGGCTCGGGGGTGTCTTTATTGACACAGAGTGGGCTCTCACAGCGGCTCACTG	308
Db	249	CAGCCTGGCTCGGGGGTGTCTTTATTGACACAGAGTGGGCTCTCACAGCGGCTCACTG	308
Qy	309	CAGCGGCAG 317	
Db	309	CAGCGGCAG 317	

RESULT 15
AC130188/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC130188 176647 bp DNA linear HTG 14-NOV-2002
Papio anubis clone RP41-421P3, WORKING DRAFT SEQUENCE, 12 ordered pieces.
AC130188 2 GI:24960890
HTG; HTGS_PHASE2; HTGS_DRAFT.
Papio anubis (olive baboon)
Papio anubis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.
1 (bases 1 to 176647)
Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 176647)
Green, E.D.
Direct Submission
Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 176647)
Green, E.D.
Direct Submission
Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
On Nov 14, 2002 this sequence version replaced gi:22138439.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: dsr
Center Clone name: 421P03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171127 bases at least Q40
Consensus quality: 173424 bases at least Q30
Consensus quality: 174803 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 175547; sum-of-contigs
Quality coverage: 8.76x in Q20 bases; agarose-fp
Quality coverage: 9.39x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1
* 6527: gap of unknown length
* 6627: contig of 6526 bp in length
* 18812: contig of 12186 bp in length
* 18912: gap of unknown length
* 18913: contig of 42868 bp in length
* 61781: gap of unknown length
* 61881: gap of 35188 bp in length
* 97069: gap of unknown length
* 97169: contig of 28038 bp in length
* 125207: gap of unknown length
* 125306: gap of 1715 bp in length
* 127021: contig of 1715 bp in length
* 127121: gap of unknown length
* 129387: contig of 2266 bp in length
* 129487: gap of unknown length
* 132302: contig of 2815 bp in length
* 132402: gap of unknown length
* 135735: contig of 3333 bp in length
* 135835: gap of unknown length
* 143558: contig of 7723 bp in length
* 143559: gap of unknown length
* 170182: contig of 26524 bp in length
* 170183: gap of unknown length
* 170283: contig of 6365 bp in length.

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 26.9%; Score 294; DB 2; Length 176647;
Best Local Similarity 91.3%; Pred. No. 1.9e-56;

	Matches	347;	Conservative	0;	Mismatches	25;	Indels	8;	Gaps	3;
QY	702	TGCTGTCAGGGTGATTC	TGCGGGGCCCTTGGTGTGTGGGGGAGTCCCTTCAAGTCTGGT							761
Db	72787	TCCTTTCCAGGGTGATTC	TGCGGGGCCCTTGGTGTGTGGGGGAGTCCCTTCAAGTCTGGT							72728
QY	762	GTCTGGGGGTCTGTGTGGGGCCCTT	TGGACAAGATGGCATCCCTGGAGTCTACACTATAT							821
Db	72727	GTCTGGGGGTCTGTGTGGGGCCCTT	TGGACAAGATGGCATCCCTGGAGTCTACACTATAT							72668
QY	822	TTGCAAGTATGTGACCTGGATCCGGATGATCATGAGGACAACCTGCTGTTCTCCAC								881
Db	72667	TTGCAAGTATGTGACCTGGATCCGGATGATCATGAGGACAACCTGCTGTTCTCCAC								72608
QY	882	CTCCACCCCAACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCCTC								941
Db	72607	CTCCATCC-----CTTAGCTTGGGTACCACTCTGACCCCTCAGAGCACCAATATCTCCTC								72554
QY	942	CATCACTTCCCTAGCTCCACTCTTGTGGCTGGAACTTTCTTGGAAC-TTTAACTCCT								1000
Db	72553	CATCGCTTTTCTGGCTCCATTCCTTGTGGCTGGAACTTTTGGAACTTTTAACTCCT								72494
QY	1001	GCAGCCCTTCTTAAGACCCACGAGC-GGGGTGAGAGAGTGTGCAATAGTCTGGAATAA								1059
Db	72493	GCAGCCCTTCTTAAGACCCACGAGC-GGGGTGAGAGAGTGTGCAATAGTCTGGAATAA								72434
QY	1060	TATAAATGAAGAGGGGCAA								1079
Db	72433	TATAAATGAAGAGGGGCAA								72414

Search completed: June 20, 2004, 03:46:14
Job time : 4628 secs

OM nucleic - nucleic search, using sw model

Run on: June 19, 2004, 23:53:03 ; Search time 508 Seconds
(without alignments)
9123.591 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: Geneseqn1980a:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result		Query		SUMMARIES		Description	
No.	Score	Match	Length	DB	ID		
1	1091	100.0	1091	3	AAA77671	Aaa77671 Human PRO	
2	1091	100.0	1091	3	AAC58114	Aac58114 Human PRO	
3	1091	100.0	1091	3	AAA37075	Aaa37075 Human PRO	
4	1091	100.0	1091	4	AAF54341	Aaf54341 DNA encod	
5	1091	100.0	1091	8	ACD68378	Acd68378 Novel hum	
6	1091	100.0	1091	8	ACH04480	Ach04480 Human cdn	
7	1091	100.0	1091	8	ACD68024	Acd68024 Novel hum	
8	1091	100.0	1091	9	ADC18062	Adc18062 Human PRO	
9	1091	100.0	1091	9	ADD70708	Add70708 Human CDN	
10	1091	100.0	1091	9	ADD39785	Add39785 Human CDN	
11	1091	100.0	1091	9	ADD70231	Add70231 Human CDN	
12	1091	100.0	1091	9	ADD38352	Add38352 Human CDN	
13	1091	100.0	1091	9	ADD39308	Add39308 Human CDN	
14	1091	100.0	1091	9	ADD38831	Add38831 Human CDN	
15	1091	100.0	1091	9	ADD40262	Add40262 Human CDN	
16	1091	100.0	1091	9	ADE50483	Ade50483 Human CDN	
17	1091	100.0	1091	9	ADE20095	Ade20095 Human CDN	
18	1091	100.0	1091	9	ADE50006	Ade50006 Human CDN	
19	1091	100.0	1091	9	ADE21564	Ade21564 Human CDN	
20	968	88.7	1185	4	AAH98653	Aah98653 Human EST	
21	746.6	68.4	747	7	AAL59914	Aal59914 Human kal	
22	727	66.6	945	6	ABK94900	Abk94900 Human nov	
23	506.4	46.4	840	2	AAx80906	Aax80906 Human cdn	

24	493.2	45.2	619	6	ABK30369	Abk30369 Human G-p
25	373.2	34.2	11820	3	AAA95944	Aaa95944 Human KLK
26	297	27.2	405	5	AAF65654	Aaf65654 Novel hum
27	219.4	20.1	1146	2	AAV84589	Aav84589 Human sec
28	219.4	20.1	1146	4	ABA83430	Aba83430 Human sec
29	219.4	20.1	1146	8	ACH04931	Ach04931 Novel hum
30	219.4	20.1	1146	8	ACD44741	Acd44741 Human cdn
31	217.8	20.0	1166	4	AAD14841	Aad14841 Human PSI
32	217.6	19.9	1105	2	AZ22638	Az22638 CASB12 nu
33	217.6	19.9	1158	2	AZ22639	Az22639 CASB12 de
34	217.6	19.9	1186	6	ABK92131	Abk92131 Prostata
35	217.6	19.9	1186	7	ABX76468	Abx76468 Lung canc
36	217.6	19.9	1186	9	ADB80595	Adb80595 Ovarian c
37	217.6	19.9	1186	9	ADB75387	Adb75387 Prostata
38	217.6	19.9	1192	4	AAD14842	Aad14842 Human PSI
39	217.6	19.9	1204	3	AAA37072	Aaa37072 Human PRO
40	217.6	19.9	1204	4	AAF54320	Aaf54320 DNA encod
41	217.6	19.9	1204	4	AA521496	Aa521496 Human CDN
42	217.6	19.9	1204	6	ABK33628	Abk33628 cDNA enco
43	217.6	19.9	1204	6	ABL88175	Ab188175 Human PRO
44	217.6	19.9	1204	6	ABL95664	Ab195664 Human ang
45	217.6	19.9	1204	7	ACD24105	Acd24105 Novel hum

ALIGNMENTS

RESULT 1

AAA77671

ID AAA77671 standard; cDNA; 1091 BP.

XX

AC AAA77671;

XX

DT 07-NOV-2000 (first entry)

XX

DE Human PRO1303 cDNA sequence SEQ ID NO:202.

XX

KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;

KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;

KW angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;

KW cytostatic; Gene therapy; vaccine; ss.

OS Homo sapiens.

XX

PN WO200032221-A2.

XX

PD 08-JUN-2000.

XX

PF 30-NOV-1999; 99WO-US028313.

XX

PR 01-DEC-1998; 98WO-US025108.

PR 16-DEC-1998; 98US-0112850P.

PR 12-JAN-1999; 99US-0115554P.

PR 08-MAR-1999; 99WO-US005028.

PR 12-MAR-1999; 99US-0123957P.

PR 28-APR-1999; 99US-013445P.

PR 14-MAY-1999; 99US-0134287P.

PR 02-JUN-1999; 99WO-US012252.

PR 23-JUN-1999; 99US-0141037P.

PR 20-JUL-1999; 99US-0144758P.

PR 26-JUL-1999; 99US-0145698P.

PR 01-SEP-1999; 99WO-US020111.

PR 08-SEP-1999; 99WO-US020594.

PR 13-SEP-1999; 99WO-US020944.

PR 15-SEP-1999; 99WO-US021090.

PR 05-OCT-1999; 99WO-US021547.

PR 29-OCT-1999; 99US-0162506P.

(GETH) GENENTECH INC.

PA Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ;

XX Goddard A, Godowski PU, Gurney AL, Klein RD, Kuo SS, Paoni NF;

PI

[illegible]

241	QY	GA	GGGCA	CCAGCCTGCGCTCGGGGGTGTCTTTATTGACACACAGTGGGTCTCACAGC	300
		DB			
241	DB	GA	GGGCA	CCAGCCTGCGCTCGGGGGTGTCTTTATTGACACACAGTGGGTCTCACAGC	300
		QY			
301	QY	GCT	CACTG	CAGCGCAGAGTACTGGGTGCGCTTGGGGGAACACAGCCTCAGCCAGTTC	360
		DB			
301	DB	GCT	CACTG	CAGCGCAGAGTACTGGGTGCGCTTGGGGGAACACAGCCTCAGCCAGTTC	360
		QY			
361	QY	GACT	GGAC	CCAGCAGATCCGGCAGACGGCTTCTCTGTGACCCATCCGGTACTCTGGGA	420
		DB			
361	DB	GACT	GGAC	CCAGCAGATCCGGCAGACGGCTTCTCTGTGACCCATCCGGTACTCTGGGA	420
		QY			
421	QY	GCT	TCGAC	AGCCACGACACGACTCCGGTGTGCGGCTGCGCTGCGCTGCGCTGCGCGTA	480
		DB			
421	DB	GCT	TCGAC	AGCCACGACACGACTCCGGTGTGCGGCTGCGCTGCGCTGCGCTGCGCGTA	480
		QY			
481	QY	AC	CAGCAGCGTTCA	ACCCCTGCCCCTGCCCAATGACTGTGCAACCGCTGGCACCCGAGTGC	540
		DB			
481	DB	AC	CAGCAGCGTTCA	ACCCCTGCCCCTGCCCAATGACTGTGCAACCGCTGGCACCCGAGTGC	540
		QY			
541	QY	CAC	GTCTC	CAGGCTGGGGCATACCAACACCGAACCCATTCCTCGGATCTGCTCCAG	600
		DB			
541	DB	CAC	GTCTC	CAGGCTGGGGCATACCAACACCGAACCCATTCCTCGGATCTGCTCCAG	600
		QY			
601	QY	TGC	CTCA	ACCTCTCCATCGTCTCCCATGCCACTGCGCTGCTGTGTATCCCGGAGATC	660
		DB			
601	DB	TGC	CTCA	ACCTCTCCATCGTCTCCCATGCCACTGCGCTGCTGTGTATCCCGGAGATC	660
		QY			
661	QY	AC	GACCA	CATGCTGTGTCAGGCGCGCTCCCGGGCAGGATGCTGCCAGGTGATCT	720
		DB			
661	DB	AC	GACCA	CATGCTGTGTCAGGCGCGCTCCCGGGCAGGATGCTGCCAGGTGATCT	720
		QY			
721	QY	GG	GGGCCCC	CTGTGTGTGGGGAGTCTTCAAGTCTGTGTCTGGGGTCTGTGGGG	780
		DB			
721	DB	GG	GGGCCCC	CTGTGTGTGGGGAGTCTTCAAGTCTGTGTCTGGGGTCTGTGGGG	780
		QY			
781	QY	CC	CTGTG	GAACAAGTGGCATCCCTGGAGTCTACACCTATATTTCGAAGTATGTGGACTGG	840
		DB			
781	DB	CC	CTGTG	GAACAAGTGGCATCCCTGGAGTCTACACCTATATTTCGAAGTATGTGGACTGG	840
		QY			
841	QY	AT	CCGGAT	GATCATGAGGAACAATGACCTGTTCCTCCACTGACCCCAACCCCTTAA	900
		DB			
841	DB	AT	CCGGAT	GATCATGAGGAACAATGACCTGTTCCTCCACTGACCCCAACCCCTTAA	900
		QY			
901	QY	CTT	GGGTAC	CCCCCTGCGCCCTCAGAGCACCAATATCTCCTCATCACTTCCTCAGTCC	960
		DB			
901	DB	CTT	GGGTAC	CCCCCTGCGCCCTCAGAGCACCAATATCTCCTCATCACTTCCTCAGTCC	960
		QY			
961	QY	ACT	CTTGT	GGCTGGGAACCTTTTGGAACTTTAACTCCTGCCAGCCCTTCTAAGACCA	1020
		DB			
961	DB	ACT	CTTGT	GGCTGGGAACCTTTTGGAACTTTAACTCCTGCCAGCCCTTCTAAGACCA	1020
		QY			
1021	QY	CG	AGCGGGT	GAGAGAAGTGTGCAATAGTCTGGATTAATATAATGAAGAGGGGCAAA	1080
		DB			
1021	DB	CG	AGCGGGT	GAGAGAAGTGTGCAATAGTCTGGATTAATATAATGAAGAGGGGCAAA	1080
		QY			
1081	QY	AAAA	AAAAAAA	1091	
		DB			
1081	DB	AAAA	AAAAAAA	1091	
		QY			
1081	QY	AAAA	AAAAAAA	1091	
		DB			
1081	DB	AAAA	AAAAAAA	1091	

RESULT 4
AAF54341
ID AAF54341 standard: DNA; 1091 BP.

XX AAF54341;

XX
DT 02-APR-2001 (first entry)

XX DE RNA encoding protein of the invention #54.

XX
continued. transmembrane: gene therapy: ss:

Secreted; transmembrane; gene therapy; ss.

XX
KM
secreted; transmembrane; gene

OS	Unidentified.
XX	WO200078961-A1.
PN	
XX	28-DEC-2000.
PD	
XX	18-FEB-2000; 200OWO-US004342.
PF	
XX	23-JUN-1999; 99US-0141037P.
PR	20-JUL-1999; 99US-0144758P.
PR	26-JUL-1999; 99US-0145698P.
PR	01-SEP-1999; 99WO-US020111.
PR	29-OCT-1999; 99US-0162506P.
PR	30-NOV-1999; 99WO-US028313.
PR	02-DEC-1999; 99WO-US028551.
PR	16-DEC-1999; 99WO-US030095.
PR	05-JAN-2000; 200OWO-US000219.
PR	08-JAN-2000; 200OWO-US000376.
XX	(GETH) GENENTECH INC.
PA	
XX	Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI	Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI	Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI	Williams PM, Wood WI;
XX	WPI; 2001-071395/08.
DR	
XX	Secreted and transmembrane proteins and nucleic acids designated PRO,
XX	useful as hybridization probes, in chromosome and gene mapping and gene
PT	therapy.
PT	
XX	Claim 2; Fig 107; 787pp; English.
PS	
XX	The present invention relates to secreted and transmembrane proteins.
CC	These proteins and the DNA encoding them may be used as hybridization
CC	probes, in chromosome and gene mapping and in the generation of anti-
CC	sense RNA and DNA. They may also be used to generate either
CC	transgenic animals or knockout animals which are in turn useful for
CC	development and screening of therapeutically useful reagents. The nucleic
CC	acids may also be used in gene therapy
XX	
SQ	Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;
Query Match 100.0%; Score 1091; DB 4; Length 1091;	
Best Local Similarity 100.0%; Pred. No. 1.5e-247;	
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CAAGCAGGTATCCCGTGGTGACCTCAAGAGAAGCAGAGGGGCAGAGGTGGGGGGC 60
DB	1 CAAGCAGGTATCCCGTGGTGACCTCAAGAGAAGCAGAGGGGCAGAGGTGGGGGGC 60
QY	61 ACAGGGAAGGTGACCTCTGAGATTCCCCTTTCCCCCAGACTTCGAAAGTGACCACC 120
DB	61 ACAGGGAAGGTGACCTCTGAGATTCCCCTTTCCCCCAGACTTCGAAAGTGACCACC 120
QY	121 ATGGGGCTCAGCATCTTTTTGCTCCTGTGTCTTGGGCTCAGCCAGGCAGCACACCG 180
DB	121 ATGGGGCTCAGCATCTTTTTGCTCCTGTGTCTTGGGCTCAGCCAGGCAGCACACCG 180
QY	181 AGATTTTCAATGGCACTGAGTGTGGGCGTAACACAGCCGTGGCAGTGGGGCTGTTT 240
DB	181 AGATTTTCAATGGCACTGAGTGTGGGCGTAACCTACAGCCGTGGCAGTGGGGCTGTTT 240
QY	241 GAGGGCACAGCTGGCGTCGGGGGTGTCCCTTATTGACCACAGGTGGGTCTCTACAGCG 300
DB	241 GAGGGCACAGCTGGCGTCGGGGGTGTCCCTTATTGACCACAGGTGGGTCTCTACAGCG 300
QY	301 GCTCACTGACGGCAGCAGTACTGGGTGCGCTGGGGGAACAACAGCCTCAGCCAGCTC 360
DB	301 GCTCACTGACGGCAGCAGTACTGGGTGCGCTGGGGGAACAACAGCCTCAGCCAGCTC 360
QY	361 GACTGGACCGAGCAGATCCGGCACAGCGGTCTCTGTGACCCATCCCGGCTAACCTGGGA 420

[illegible]

06-OCT-1998;	98US-0103443P;
07-OCT-1998;	98US-0103314P;
07-OCT-1998;	98US-0103315P;
07-OCT-1998;	98US-0103328P;
07-OCT-1998;	98US-0103395P;
07-OCT-1998;	98US-0103396P;
07-OCT-1998;	98US-0103401P;
08-OCT-1998;	98US-0103633P;
08-OCT-1998;	98US-0103678P;
08-OCT-1998;	98US-0103679P;
08-OCT-1998;	98US-0103711P;
14-OCT-1998;	98US-0104257P;
20-OCT-1998;	98US-0104987P;
20-OCT-1998;	98US-0105000P;
20-OCT-1998;	98US-0105002P;
21-OCT-1998;	98US-0105104P;
22-OCT-1998;	98US-0105169P;
22-OCT-1998;	98US-0105266P;
26-OCT-1998;	98US-0105693P;
26-OCT-1998;	98US-0105694P;
27-OCT-1998;	98US-0105807P;
27-OCT-1998;	98US-0105881P;
27-OCT-1998;	98US-0105882P;
27-OCT-1998;	98US-0106062P;
28-OCT-1998;	98US-0106023P;
28-OCT-1998;	98US-0106029P;
28-OCT-1998;	98US-0106030P;
28-OCT-1998;	98US-0106032P;
28-OCT-1998;	98US-0106033P;
28-OCT-1998;	98US-0106178P;
29-OCT-1998;	98US-0106248P;
29-OCT-1998;	98US-0106384P;
29-OCT-1998;	98US-0108500P;
30-OCT-1998;	98US-0106464P;
03-NOV-1998;	98US-0106956P;
03-NOV-1998;	98US-0106902P;
03-NOV-1998;	98US-0106905P;
03-NOV-1998;	98US-0106919P;
03-NOV-1998;	98US-0106932P;
03-NOV-1998;	98US-0106934P;
10-NOV-1998;	98US-0107783P;
17-NOV-1998;	98US-0108775P;
17-NOV-1998;	98US-0108787P;
17-NOV-1998;	98US-0108788P;
17-NOV-1998;	98US-0108801P;
17-NOV-1998;	98US-0108802P;
17-NOV-1998;	98US-0108806P;
17-NOV-1998;	98US-0108807P;
17-NOV-1998;	98US-0108867P;
17-NOV-1998;	98US-0108925P;
17-NOV-1998;	98US-0108848P;
18-NOV-1998;	98US-0108850P;
18-NOV-1998;	98US-0108851P;
18-NOV-1998;	98US-0108852P;
18-NOV-1998;	98US-0108904P;
22-DEC-1998;	98US-0113296P;
30-DEC-1998;	98US-0114223P;
05-JAN-1999;	98WO-US000106;
16-APR-1999;	98US-0129674P;
23-JUN-1999;	99US-0141037P;
20-JUL-1999;	99US-0144758P;
26-SEP-1999;	99US-0145699P;
01-SEP-1999;	99WO-US020111;
15-SEP-1999;	99WO-US021194;
23-OCT-1999;	99US-0162506P;
30-OCT-1999;	99WO-US028313;
02-DEC-1999;	99WO-US028551;
15-DEC-1999;	99WO-US030095;
05-JAN-2000;	2000WO-US000219;
06-JAN-2000;	2000WO-US000376;

[illegible]

[illegible]

[illegible]

Mon Jun 21 09:00:25 2004

us-10-015-385a-193.rng

```
Db      721  |||||GGGGCCCCCTGGTGTGTGGGGAGTCCCTTCAAGGTCCTGGTGTCTGGGGTCTGTGGGG 780
Qy      781  CCCTGTGGACAAGATGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840
Db      781  CCCTGTGGACAAGATGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840
Qy      841  ATCCGATGATCATGAGGACAACTGACCTGTTTCTCCACCTCCACCCCTTAA 900
Db      841  ATCCGATGATCATGAGGACAACTGACCTGTTTCTCCACCTCCACCCCTTAA 900
Qy      901  CTTGGGTACCCCTCTGGCCCTCAGAGCACAATATCTCTCCATCTCCCTAGCTCC 960
Db      901  CTTGGGTACCCCTCTGGCCCTCAGAGCACAATATCTCTCCATCTCCCTAGCTCC 960
Qy      961  ACTCTGTGGCTGGGAACCTTCTTGAACCTTTAACTCTCCGAGCCCTTCTAAGACCCA 1020
Db      961  ACTCTGTGGCTGGGAACCTTCTTGAACCTTTAACTCTCCGAGCCCTTCTAAGACCCA 1020
Qy      1021 CGAGCGGGTGGAGAGAGTGTGCAATAGTCTGGAAATAATAATGAAGAGGGGCAAA 1080
Db      1021 CGAGCGGGTGGAGAGAGTGTGCAATAGTCTGGAAATAATAATGAAGAGGGGCAAA 1080
Qy      1081 AAAAAAAAAA 1091
Db      1081 AAAAAAAAAA 1091

RESULT 8
ADCI18062
ID ADCI18062 standard; cDNA; 1091 BP.
XX
AC ADCI18062;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human PRO polynucleotide #54.
XX
KW Human; PRO; gene; ss; protein electrophoresis; chromosome mapping;
KW gene mapping; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003064925-A1.
XX
XX
PD 03-APR-2003.
XX
XX
PF 10-DEC-2001; 2001US-00013907.
XX
XX
XX 01-SEP-1998; 98US-0098716P.
XX 01-SEP-1998; 98US-0098723P.
XX 01-SEP-1998; 98US-0098749P.
XX 01-SEP-1998; 98US-0098750P.
XX 01-SEP-1998; 98US-0098803P.
XX 02-SEP-1998; 98US-0098821P.
XX 02-SEP-1998; 98US-0098843P.
XX 02-SEP-1998; 98US-0099536P.
XX 09-SEP-1998; 98US-0099596P.
XX 09-SEP-1998; 98US-0099598P.
XX 09-SEP-1998; 98US-0099602P.
XX 09-SEP-1998; 98US-0099642P.
XX 09-SEP-1998; 98US-0099741P.
XX 10-SEP-1998; 98US-0099754P.
XX 10-SEP-1998; 98US-0099792P.
XX 10-SEP-1998; 98US-0099808P.
XX 10-SEP-1998; 98US-0099812P.
XX 10-SEP-1998; 98US-0099815P.
XX 10-SEP-1998; 98US-0099816P.
XX 15-SEP-1998; 98US-0100385P.
XX 15-SEP-1998; 98US-0100389P.
XX 15-SEP-1998; 98US-0100390P.
XX 15-SEP-1998; 98US-0100584P.
XX 16-SEP-1998; 98US-0100584P.
XX 16-SEP-1998; 98US-0100661P.
XX 16-SEP-1998; 98US-0100662P.
XX 16-SEP-1998; 98US-0100664P.
XX 16-SEP-1998; 98US-0100683P.
XX 17-SEP-1998; 98US-0100684P.
XX 17-SEP-1998; 98US-0100710P.
XX 17-SEP-1998; 98US-0100711P.
XX 17-SEP-1998; 98US-0100919P.
XX 17-SEP-1998; 98US-0100930P.
XX 17-SEP-1998; 98US-0100948P.
XX 18-SEP-1998; 98US-0100849P.
XX 18-SEP-1998; 98US-0101014P.
XX 18-SEP-1998; 98US-0101068P.
XX 18-SEP-1998; 98US-0101071P.
XX 22-SEP-1998; 98US-0101279P.
XX 22-SEP-1998; 98US-0101471P.
XX 23-SEP-1998; 98US-0101472P.
XX 23-SEP-1998; 98US-0101474P.
XX 23-SEP-1998; 98US-0101475P.
XX 23-SEP-1998; 98US-0101476P.
XX 23-SEP-1998; 98US-0101477P.
XX 24-SEP-1998; 98US-0101479P.
XX 24-SEP-1998; 98US-0101738P.
XX 24-SEP-1998; 98US-0101741P.
XX 24-SEP-1998; 98US-0101743P.
XX 24-SEP-1998; 98US-0101915P.
XX 24-SEP-1998; 98US-0101916P.
XX 24-SEP-1998; 98US-010207P.
XX 29-SEP-1998; 98US-0102240P.
XX 29-SEP-1998; 98US-0102307P.
XX 29-SEP-1998; 98US-0102330P.
XX 29-SEP-1998; 98US-0102331P.
XX 30-SEP-1998; 98US-0102484P.
XX 30-SEP-1998; 98US-0102487P.
XX 30-SEP-1998; 98US-0102570P.
XX 30-SEP-1998; 98US-0102571P.
XX 01-OCT-1998; 98US-0102684P.
XX 01-OCT-1998; 98US-0102687P.
XX 02-OCT-1998; 98US-0102965P.
XX 06-OCT-1998; 98US-0103258P.
XX 06-OCT-1998; 98US-0103449P.
XX 07-OCT-1998; 98US-0103314P.
XX 07-OCT-1998; 98US-0103315P.
XX 07-OCT-1998; 98US-0103328P.
XX 07-OCT-1998; 98US-0103395P.
XX 07-OCT-1998; 98US-0103396P.
XX 07-OCT-1998; 98US-0103401P.
XX 08-OCT-1998; 98US-0103633P.
XX 08-OCT-1998; 98US-0103678P.
XX 08-OCT-1998; 98US-0103679P.
XX 14-OCT-1998; 98US-0103711P.
XX 20-OCT-1998; 98US-0104257P.
XX 20-OCT-1998; 98US-0104987P.
XX 20-OCT-1998; 98US-0105000P.
XX 21-OCT-1998; 98US-0105104P.
XX 22-OCT-1998; 98US-0105169P.
XX 22-OCT-1998; 98US-0105266P.
XX 26-OCT-1998; 98US-0105693P.
XX 26-OCT-1998; 98US-0105807P.
XX 27-OCT-1998; 98US-0105881P.
XX 27-OCT-1998; 98US-0105882P.
XX 28-OCT-1998; 98US-0106023P.
XX 28-OCT-1998; 98US-0106030P.
XX 28-OCT-1998; 98US-0106032P.
XX 28-OCT-1998; 98US-0106033P.
XX 28-OCT-1998; 98US-0106178P.
XX 29-OCT-1998; 98US-0106248P.
XX 29-OCT-1998; 98US-0106384P.
```


16-SEP-1998; 98US-0100664P

PR	17-SEP-1998;	98US-0100683P;
PR	17-SEP-1998;	98US-0100684P;
PR	17-SEP-1998;	98US-0100710P;
PR	17-SEP-1998;	98US-0100711P;
PR	17-SEP-1998;	98US-0100919P;
PR	17-SEP-1998;	98US-0100930P;
PR	17-SEP-1998;	98US-0100848P;
PR	18-SEP-1998;	98US-0100849P;
PR	18-SEP-1998;	98US-0101014P;
PR	18-SEP-1998;	98US-0101068P;
PR	18-SEP-1998;	98US-0101071P;
PR	22-SEP-1998;	98US-0101279P;
PR	23-SEP-1998;	98US-0101471P;
PR	23-SEP-1998;	98US-0101472P;
PR	23-SEP-1998;	98US-0101474P;
PR	23-SEP-1998;	98US-0101475P;
PR	23-SEP-1998;	98US-0101476P;
PR	23-SEP-1998;	98US-0101477P;
PR	23-SEP-1998;	98US-0101479P;
PR	23-SEP-1998;	98US-0101738P;
PR	24-SEP-1998;	98US-0101741P;
PR	24-SEP-1998;	98US-0101743P;
PR	24-SEP-1998;	98US-0101915P;
PR	24-SEP-1998;	98US-0101916P;
PR	29-SEP-1998;	98US-0102207P;
PR	29-SEP-1998;	98US-0102240P;
PR	29-SEP-1998;	98US-0102307P;
PR	29-SEP-1998;	98US-0102330P;
PR	29-SEP-1998;	98US-0102331P;
PR	30-SEP-1998;	98US-0102484P;
PR	30-SEP-1998;	98US-0102487P;
PR	30-SEP-1998;	98US-0102570P;
PR	30-SEP-1998;	98US-0102571P;
PR	01-OCT-1998;	98US-0102684P;
PR	01-OCT-1998;	98US-0102687P;
PR	01-OCT-1998;	98US-0102655P;
PR	06-OCT-1998;	98US-0103258P;
PR	06-OCT-1998;	98US-0103449P;
PR	07-OCT-1998;	98US-0103314P;
PR	07-OCT-1998;	98US-0103315P;
PR	07-OCT-1998;	98US-0103328P;
PR	07-OCT-1998;	98US-0103335P;
PR	07-OCT-1998;	98US-0103339P;
PR	08-OCT-1998;	98US-0103401P;
PR	08-OCT-1998;	98US-0103633P;
PR	08-OCT-1998;	98US-0103678P;
PR	08-OCT-1998;	98US-0103679P;
PR	08-OCT-1998;	98US-0103711P;
PR	14-OCT-1998;	98US-0104257P;
PR	20-OCT-1998;	98US-0104987P;
PR	20-OCT-1998;	98US-0105002P;
PR	20-OCT-1998;	98US-0105002P;
PR	21-OCT-1998;	98US-0105104P;
PR	21-OCT-1998;	98US-0105169P;
PR	22-OCT-1998;	98US-0105266P;
PR	26-OCT-1998;	98US-0105393P;
PR	26-OCT-1998;	98US-0105694P;
PR	27-OCT-1998;	98US-0105881P;
PR	27-OCT-1998;	98US-0105882P;
PR	27-OCT-1998;	98US-0105882P;
PR	28-OCT-1998;	98US-0106023P;
PR	28-OCT-1998;	98US-0106030P;
PR	28-OCT-1998;	98US-0106032P;
PR	28-OCT-1998;	98US-0106033P;
PR	28-OCT-1998;	98US-0106178P;
PR	29-OCT-1998;	98US-0106248P;
PR	29-OCT-1998;	98US-0106384P;
PR	29-OCT-1998;	98US-0108500P;
PR	30-OCT-1998;	98US-0108464P;
PR	30-NOV-1998;	98US-0106856P;
PR	03-NOV-1998;	98US-0106902P;

WFI; 2003-755122/71.
P-PSDB; ADD39786.

New secreted and transmembrane PRO polypeptides useful for treating Crohn's disease, diabetes mellitus, hyper- or hypo-insulinemia, sports injuries and arthritis.

Claim 2: SEQ ID NO 193: 557pp; English.

The invention relates to an isolated PRO polypeptide (secreted or transmembrane protein) having at least 80% amino acid sequence identity to an amino acid sequence chosen from 123 fully defined sequences as given in the specification (including their extracellular domains either with or without their associated signal peptides). Also include are the nucleotide (NA) sequences encoding PRO, a vector comprising the PRO NA, a host cell comprising the vector, producing PRO, a chimeric molecule comprising PRO fused to a heterologous amino acid sequence, and an anti-PRO antibody. Pro is useful as molecular weight markers for protein electrophoresis and also for chromosome identification. PRO is also useful for tissue typing. PRO and PRO NA are useful as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA. PRO NA is useful for generating transgenic animals or knock-out animals which are useful in development and screening useful reagents. PRO NA is also useful in gene therapy. PRO1244, PRO1286 and PRO1303 polypeptides are useful for treating cancerous tumours. PRO1250, PRO1418 and PRO1410 polypeptides are useful for suppressing immune response. PRO1246 polypeptide is useful for treating cardiac insufficiency disorders. PRO1245 polypeptide is also useful for treating tumours. PRO1246 and PRO1361 polypeptide are useful for stimulating calcium flux in human umbilical vein endothelial cells. PRO1265, PRO1250 and PRO1474 polypeptides are useful for treating bone and/or cartilage disorders (e.g., arthritis) and wound healing. PRO1130, PRO1275 and PRO1418 polypeptides are useful for treating diabetes in skeletal muscle cells and obesity. PRO1265, PRO1244 and PRO1382 polypeptides are useful for treating Berger disease or other nephropathies associated with Schnlein-Henoch purpura, coeliac disease, dermatitis, herpeticiformis or Crohn's disease. PRO1478, PRO1265, PRO1412, PRO1279, PRO1304, PRO1306, PRO1418, PRO1410 and PRO1575 are useful in treating thalassaemias. The present invention also encompasses a PRO protein of the invention.

Sequence	1091 BP;	221 A;	335 C;	313 G;	222 T;	0 U;	0 Other;
Query Match	100.0%;	Score 1091;	DB 9;	Length 1091;			
Best Local Similarity	100.0%;	Pred. No. 1.5e-247;					
Matches 1091;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
1	CAAGCAGGTCATCCCTTGGTGACCTTTAAAGAGAACGACAGAGGCACAGGTGGGGGCGC	60					
1	CAAGCAGGTCATCCCTTGGTGACCTTTCAAGAGAGAACGACAGAGGCACAGGTGGGGGCGC	60					
61	ACAGGGAAAGGGTGACCTCTGAGATCCCTTTTCCCCAGACTTTTGGAAAGTGACCCACCC	120					
61	ACAGGGAAAGGGTGACCTCTGAGATCCCTTTTCCCCAGACTTTTGGAAAGTGACCCACCC	120					
121	ATGGGGCTCAGCATCTTTTGTGCTCTGTGTGTTCTTGGGCTCAGCAGGACGCCACACGG	180					
121	ATGGGGCTCAGCATCTTTTGTGCTCTGTGTGTTCTTGGGCTCAGCAGGACGCCACACGG	180					
181	AAGATTTTCAATGGCATCTGAGTGTGGGCGTAACTACAGCCGTGCAGGTGGGGCTGTTT	240					
181	AAGATTTTCAATGGCATCTGAGTGTGGGCGTAACTACAGCCGTGCAGGTGGGGCTGTTT	240					
241	GAGGGCACCAAGCTGGCGCTCGCGGGGTGCTCTTATTGACACAGGTGGGTCTCTCACAGCG	300					
241	GAGGGCACCAAGCTGGCGCTCGCGGGGTGCTCTTATTGACCAAGGTGGGTCTCTCACAGCG	300					
301	GCTCACTGCAGCGGCACAGGTACTTGGGTGCGCTGTGGGGGAACACAGCTCAGCCAGCTC	360					
301	GCTCACTGCAGCGGCACAGGTACTTGGGTGCGCTGTGGGGGAACACAGCTCAGCCAGCTC	360					
361	GACTTGGACCGAGCAGATCCGGGCACAGCGGCTTCTGTGTGACCCATCCCGGTACCTGGGA	420					

(GETH) GENENTECH INC.

Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;
 Xi
 Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PPI
 Pan J, Paonl NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
 PPI
 Williams PM, Wood WI;
 PPI

PK PR

Mon Jun 21 09:00:25 2004

Db	601	TGCTCAACCTCTCCATCGCTCCCATGCACTGCCATGGTGTATGCCGGGAGAAATC	660	PR	10-SEP-1998;	98US-0099754P.
Qy	661	ACGAGCAACATGTTGTGTGAGCGGGCTCCCGGGCAGGATGCCTGCCAGGTTGATTCT	720	PR	10-SEP-1998;	98US-0099763P.
Db	661	ACGAGCAACATGTTGTGTGAGCGGGCTCCCGGGCAGGATGCCTGCCAGGTTGATTCT	720	PR	10-SEP-1998;	98US-0099792P.
Qy	721	GGGGCCCCCTGTGTGTGGGGAGTCCCTCAAGTCTGTGTCTGTGGGGTCTGTGGGG	780	PR	10-SEP-1998;	98US-0099808P.
Db	721	GGGGCCCCCTGTGTGTGGGGAGTCCCTCAAGTCTGTGTCTGTGGGGTCTGTGGGG	780	PR	10-SEP-1998;	98US-0099812P.
Qy	781	CCCTGTGGCAAGATGCATCCCTGGAGTCTACACCTATATTTCGAAGTATGTGGACTGG	840	PR	10-SEP-1998;	98US-0099815P.
Db	781	CCCTGTGGCAAGATGCATCCCTGGAGTCTACACCTATATTTCGAAGTATGTGGACTGG	840	PR	10-SEP-1998;	98US-0099816P.
Qy	841	ATCCGGATGATCATGAGGAAACAATGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA	900	PR	10-SEP-1998;	98US-0100385P.
Db	841	ATCCGGATGATCATGAGGAAACAATGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA	900	PR	10-SEP-1998;	98US-0100388P.
Qy	901	CTTGGATACCCCTCTGGCCCTCAGAGCAACCAATATCTCTCCATCATCTTCCCTAGCTCC	960	PR	10-SEP-1998;	98US-0100642P.
Db	901	CTTGGATACCCCTCTGGCCCTCAGAGCAACCAATATCTCTCCATCATCTTCCCTAGCTCC	960	PR	10-SEP-1998;	98US-0100643P.
Qy	961	ACTCTGTGGCTGGGAACTTCTTGGAACTTTAACTCTGCCAGCCCTTCTTAAGACCCA	1020	PR	10-SEP-1998;	98US-0100662P.
Db	961	ACTCTGTGGCTGGGAACTTCTTGGAACTTTAACTCTGCCAGCCCTTCTTAAGACCCA	1020	PR	10-SEP-1998;	98US-0100664P.
Qy	1021	CGAGCGGGTGCAGAGAGTGTGCAATAGTCTGGAATAAATAAATGAAGAGGGGGGAAA	1080	PR	10-SEP-1998;	98US-0100683P.
Db	1021	CGAGCGGGTGCAGAGAGTGTGCAATAGTCTGGAATAAATAAATGAAGAGGGGGGAAA	1080	PR	10-SEP-1998;	98US-0100684P.
Qy	1081	AAAAAAAAAAAA 1091		PR	10-SEP-1998;	98US-0100711P.
Db	1081	AAAAAAAAAAAA 1091		PR	10-SEP-1998;	98US-0100712P.
RESULT 13						
AD	ADD39308	standard; cDNA; 1091 BP.		PR	10-SEP-1998;	98US-0101014P.
XX	AC			PR	10-SEP-1998;	98US-0101068P.
XX	ADD39308;			PR	10-SEP-1998;	98US-0101071P.
DT	15-JAN-2004	(first entry)		PR	10-SEP-1998;	98US-0101279P.
XX	Human cDNA encoding secreted/transmembrane protein PRO1303.			PR	10-SEP-1998;	98US-0101476P.
XX	Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;			PR	10-SEP-1998;	98US-0101477P.
KW	immune response; cardiac insufficiency disorder; calcium flux;			PR	10-SEP-1998;	98US-0101479P.
KW	umbilical vein endothelial cell; bone disorder; cartilage disorder;			PR	10-SEP-1998;	98US-0101738P.
KW	arthritis; wound healing; diabetes; skeletal muscle cells; obesity;			PR	10-SEP-1998;	98US-0101741P.
KW	Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;			PR	10-SEP-1998;	98US-0101743P.
KW	dermatitis; herpeticiformis; Crohn's disease; thalassemia.			PR	10-SEP-1998;	98US-0101915P.
OS	Homo sapiens.			PR	10-SEP-1998;	98US-0102207P.
XX	US2003096954-A1.			PR	10-SEP-1998;	98US-0102240P.
PN	22-MAY-2003.			PR	10-SEP-1998;	98US-0102307P.
PD				PR	10-SEP-1998;	98US-0102331P.
XX	07-DEC-2001; 2001US-00011671.			PR	10-SEP-1998;	98US-0102484P.
XX				PR	10-SEP-1998;	98US-0102487P.
PR	01-SEP-1998;	98US-0098716P.		PR	10-SEP-1998;	98US-0102570P.
PR	01-SEP-1998;	98US-0098723P.		PR	10-SEP-1998;	98US-0102571P.
PR	01-SEP-1998;	98US-0098749P.		PR	10-SEP-1998;	98US-0102684P.
PR	01-SEP-1998;	98US-0098750P.		PR	10-SEP-1998;	98US-0102687P.
PR	02-SEP-1998;	98US-0098803P.		PR	10-SEP-1998;	98US-0102965P.
PR	02-SEP-1998;	98US-0098821P.		PR	10-SEP-1998;	98US-0103258P.
PR	02-SEP-1998;	98US-0098843P.		PR	10-SEP-1998;	98US-0103449P.
PR	09-SEP-1998;	98US-0099536P.		PR	10-SEP-1998;	98US-0103449P.
PR	09-SEP-1998;	98US-0099596P.		PR	10-SEP-1998;	98US-0103314P.
PR	09-SEP-1998;	98US-0099598P.		PR	10-SEP-1998;	98US-0103328P.
PR	09-SEP-1998;	98US-0099602P.		PR	10-SEP-1998;	98US-0103358P.
PR	09-SEP-1998;	98US-0099642P.		PR	10-SEP-1998;	98US-0103396P.
PR	10-SEP-1998;	98US-0099741P.		PR	10-SEP-1998;	98US-0103401P.
PR				PR	10-SEP-1998;	98US-0103633P.
PR				PR	10-SEP-1998;	98US-0103678P.
PR				PR	10-SEP-1998;	98US-0103679P.
PR				PR	10-SEP-1998;	98US-0103711P.
PR				PR	10-SEP-1998;	98US-0104257P.
PR				PR	10-SEP-1998;	98US-0104987P.
PR				PR	10-SEP-1998;	98US-0105000P.
PR				PR	10-SEP-1998;	98US-0105002P.
PR				PR	10-SEP-1998;	98US-0105104P.
PR				PR	10-SEP-1998;	98US-0105169P.
PR				PR	10-SEP-1998;	98US-0105266P.
PR				PR	10-SEP-1998;	98US-0105693P.
PR				PR	10-SEP-1998;	98US-0105694P.
PR				PR	10-SEP-1998;	98US-0105807P.

[illegible]


```
Db      841 ATCCGATGATCATGAGAACAACTGACCTGTTCTCCACCTCCACCCCAACCCCTTAA 900
QY      901 CTTGGTACCCCTCTGGCCCTCAGAGCAACCAATATCTCTCCATCAGTCCCTAGCTCC 960
Db      901 CTTGGTACCCCTCTGGCCCTCAGAGCAACCAATATCTCTCCATCAGTCCCTAGCTCC 960
QY      961 ACTCTGTGGCTGGGACTCTTTGGAACTTTAACTCTGCAGCCCTTCTAAGACCCA 1020
Db      961 ACTCTGTGGCTGGGACTCTTTGGAACTTTAACTCTGCAGCCCTTCTAAGACCCA 1020
QY      1021 CGAGCGGGTGAGAGAAGTGTCATAGTCTCGAATAAATAAATGAAGAGGGGCAAA 1080
Db      1021 CGAGCGGGTGAGAGAAGTGTCATAGTCTCGAATAAATAAATGAAGAGGGGCAAA 1080
QY      1081 AAAAAAAAAA 1091
Db      1081 AAAAAAAAAA 1091

RESULT 15
ADD40262
ID   ADD40262 standard; cDNA, 1091 BP.
AC   ADD40262;
XX
XX
DT   15-JAN-2004 (first entry)
XX
DE   Human cDNA encoding secreted/transmembrane protein PRO1303.
XX
KW   Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW   immune response; cardiac insufficiency disorder; calcium flux;
KW   umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW   arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW   Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW   dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX
OS   Homo sapiens.
XX
PN   US2003082627-A1.
XX
PD   01-MAY-2003.
XX
PF   06-DEC-2001; 2001US-00006117.
XX
PR   01-SEP-1998; 98US-0098716P.
PR   01-SEP-1998; 98US-0098723P.
PR   01-SEP-1998; 98US-0098749P.
PR   01-SEP-1998; 98US-0098750P.
PR   02-SEP-1998; 98US-0098803P.
PR   02-SEP-1998; 98US-0098821P.
PR   02-SEP-1998; 98US-0098843P.
PR   02-SEP-1998; 98US-0098936P.
PR   02-SEP-1998; 98US-0098953P.
PR   02-SEP-1998; 98US-0098958P.
PR   02-SEP-1998; 98US-0098960P.
PR   02-SEP-1998; 98US-0098964P.
PR   02-SEP-1998; 98US-00989741P.
PR   02-SEP-1998; 98US-00989754P.
PR   02-SEP-1998; 98US-00989763P.
PR   02-SEP-1998; 98US-00989792P.
PR   02-SEP-1998; 98US-00989808P.
PR   02-SEP-1998; 98US-00989812P.
PR   02-SEP-1998; 98US-00989815P.
PR   02-SEP-1998; 98US-00989816P.
PR   02-SEP-1998; 98US-0100388P.
PR   02-SEP-1998; 98US-0100390P.
PR   02-SEP-1998; 98US-0100384P.
PR   02-SEP-1998; 98US-0100627P.
PR   02-SEP-1998; 98US-0100661P.
PR   02-SEP-1998; 98US-0100662P.
PR   02-SEP-1998; 98US-0100664P.
PR   02-SEP-1998; 98US-0100683P.
PR   17-SEP-1998; 98US-0100684P.
PR   17-SEP-1998; 98US-0100710P.
PR   17-SEP-1998; 98US-0100711P.
PR   17-SEP-1998; 98US-0100919P.
PR   17-SEP-1998; 98US-0100930P.
PR   18-SEP-1998; 98US-0100848P.
PR   18-SEP-1998; 98US-0100849P.
PR   18-SEP-1998; 98US-0101014P.
PR   18-SEP-1998; 98US-0101068P.
PR   18-SEP-1998; 98US-0101071P.
PR   22-SEP-1998; 98US-0101279P.
PR   23-SEP-1998; 98US-0101471P.
PR   23-SEP-1998; 98US-0101472P.
PR   23-SEP-1998; 98US-0101474P.
PR   23-SEP-1998; 98US-0101475P.
PR   23-SEP-1998; 98US-0101476P.
PR   23-SEP-1998; 98US-0101477P.
PR   23-SEP-1998; 98US-0101479P.
PR   24-SEP-1998; 98US-0101738P.
PR   24-SEP-1998; 98US-0101741P.
PR   24-SEP-1998; 98US-0101743P.
PR   24-SEP-1998; 98US-0101915P.
PR   24-SEP-1998; 98US-0101916P.
PR   29-SEP-1998; 98US-0102207P.
PR   29-SEP-1998; 98US-0102240P.
PR   29-SEP-1998; 98US-0102307P.
PR   29-SEP-1998; 98US-0102330P.
PR   29-SEP-1998; 98US-0102331P.
PR   30-SEP-1998; 98US-0102484P.
PR   30-SEP-1998; 98US-0102487P.
PR   30-SEP-1998; 98US-0102570P.
PR   30-SEP-1998; 98US-0102571P.
PR   01-OCT-1998; 98US-0102684P.
PR   01-OCT-1998; 98US-0102687P.
PR   02-OCT-1998; 98US-0102965P.
PR   06-OCT-1998; 98US-0103258P.
PR   06-OCT-1998; 98US-0103449P.
PR   07-OCT-1998; 98US-0103314P.
PR   07-OCT-1998; 98US-0103315P.
PR   07-OCT-1998; 98US-0103328P.
PR   07-OCT-1998; 98US-0103395P.
PR   07-OCT-1998; 98US-0103396P.
PR   07-OCT-1998; 98US-0103401P.
PR   08-OCT-1998; 98US-0103633P.
PR   08-OCT-1998; 98US-0103678P.
PR   08-OCT-1998; 98US-0103679P.
PR   08-OCT-1998; 98US-0103711P.
PR   14-OCT-1998; 98US-0104257P.
PR   20-OCT-1998; 98US-0104987P.
PR   20-OCT-1998; 98US-0105000P.
PR   21-OCT-1998; 98US-0105002P.
PR   21-OCT-1998; 98US-0105104P.
PR   22-OCT-1998; 98US-0105169P.
PR   22-OCT-1998; 98US-0105266P.
PR   26-OCT-1998; 98US-0105693P.
PR   26-OCT-1998; 98US-0105694P.
PR   27-OCT-1998; 98US-0105807P.
PR   27-OCT-1998; 98US-0105881P.
PR   27-OCT-1998; 98US-0105882P.
PR   27-OCT-1998; 98US-0106062P.
PR   28-OCT-1998; 98US-0106023P.
PR   28-OCT-1998; 98US-0106029P.
PR   28-OCT-1998; 98US-0106030P.
PR   28-OCT-1998; 98US-0106032P.
PR   28-OCT-1998; 98US-0106033P.
PR   28-OCT-1998; 98US-0106178P.
PR   29-OCT-1998; 98US-0106248P.
PR   29-OCT-1998; 98US-0106384P.
PR   29-OCT-1998; 98US-0108500P.
PR   30-OCT-1998; 98US-0106464P.
PR   03-NOV-1998; 98US-0106856P.
PR   03-NOV-1998; 98US-0106902P.
PR   03-NOV-1998; 98US-0106905P.
```


Db 961 ACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCCTGCAGCCCTTCTTAAGACCCA 1020
Qy 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAATGAAGGAGGGGCAAA 1080
Db 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAATGAAGGAGGGGCAAA 1080
Qy 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091

Search completed: June 20, 2004, 02:29:02
Job time : 514 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 01:58:51 ; Search time 3353 Seconds
(without alignments)
9716.564 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caacaggtatcccttgg.....agggcaaaaaaaaaa 1091

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	459	42.1	890	11 AK009217	AK009217 Mus muscu
2	443	40.6	898	13 BY709025	BY709025 BY709025
3	375	34.4	472	13 BX110362	BX110362 BX110362
4	311.8	28.6	469	9 AI394679	AI394679 tg24b03.x

c	5	305.2	28.0	676	13	BQ195790
	6	282.6	25.9	520	13	BX514743
c	7	268.2	24.7	545	9	AI324852
	8	269.4	24.6	506	9	AA028356
	9	257.2	23.6	449	9	AI322408
	10	250.4	23.0	1186	11	BC035385
	11	221.6	20.3	391	14	W20715
	12	216	19.8	264	12	BG230967
	13	208.4	19.1	1282	11	BC015551
	14	200.4	18.4	1295	11	AK009360
	15	200.4	18.4	1295	11	AK009720
	16	196	18.0	783	29	AY410899
	17	195.4	17.9	783	29	AY410900
	18	195.4	17.9	880	14	CB202840
	19	195.4	17.9	898	14	CB204935
	20	191.6	17.6	783	29	AY410898
	21	189.2	17.3	1269	11	AK009659
	22	188.4	17.3	852	14	CB587168
	23	182.8	16.8	678	12	BI763040
	24	180.6	16.6	826	14	CB574882
	25	180.4	16.5	1072	12	BM559782
	26	174.8	16.0	1074	12	BM559617
	27	174.2	16.0	841	12	BI818697
	28	174.2	16.0	853	12	EG720793
	29	169.2	15.5	754	12	BI653899
c	30	167.8	15.4	1051	12	BM547198
	31	166	15.2	730	14	CD497946
	32	166	15.2	762	14	CD497945
c	33	166	15.2	783	14	CD498690
c	34	166	15.2	785	14	CD506723
c	35	166	15.2	787	14	CD503321
c	36	166	15.2	788	14	CD506194
c	37	166	15.2	799	14	CD509135
c	38	166	15.2	800	14	CD506130
c	39	166	15.2	800	14	CD509521
	40	166	15.2	809	14	CD506193
	41	166	15.2	821	14	CD506722
	42	166	15.2	824	14	CD509134
	43	166	15.2	825	14	CD503320
	44	166	15.2	828	14	CD501456
	45	166	15.2	837	14	CD505738

ALIGNMENTS

RESULT 1
AK009217

LOCUS

DEFINITION

AK009217 890 bp mRNA linear HTC 20-SEP-2003
Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:231008B01 product:similar to KALLIKREIN 12
PRECURSOR (EC 3.4.21.-) (KALLIKREIN-LIKE PROTEIN 5) (KLK-L5) [Homo
sapiens], full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AK009217 890 bp mRNA linear HTC 20-SEP-2003
Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:231008B01 product:similar to KALLIKREIN 12
PRECURSOR (EC 3.4.21.-) (KALLIKREIN-LIKE PROTEIN 5) (KLK-L5) [Homo
sapiens], full insert sequence.

AK009217.1 GI:12843870
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalizaton and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI CGAP CLL1"
/note="Vector: pT73D-pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGCATGCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 34.4%; Score 375; DB 13; Length 472;
Best Local Similarity 99.7%; Pred. No. 6.4e-62;
Matches 375; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 100 AGACTTTGGAAGTACCCACCATGGGGCTCAGCATCTTTTGTCTCTCTGTCTTTGGG 159
DB 97 AGACTTTGGAAGTACCCACCATGGGGCTCAGCATCTTTTGTCTCTCTGTCTTTGGG 156
QY 160 CTGAGCCAGGACGACCAACCGAAGATTTTCAATGGCACTGAGTGTGGCGTAATCACAG 219
DB 157 CTGAGCCAGGACGACCAACCGAAGATTTTCAATGGCACTGAGTGTGGCGTAATCACAG 216
QY 220 CCGTGGCAGTGGGGCTGTTTGAAGGACACGAGCTGGCGTGGCGGGGTGTCCTATTGAC 279
DB 217 CCGTGGCAGTGGGGCTGTTTGAAGGACACGAGCTGGCGTGGCGGGGTGTCCTATTGAC 276
QY 280 CACAGTGGGTCTCTACAGCGGCTCACTGAGCGGACGAGTACTGGGTGGCGCTGGGG 339
DB 277 CACAGTGGGTCTCTACAGCGGCTCACTGAGCGGACGAGTACTGGGTGGCGCTGGGG 336
QY 340 GAACACAGCCTCAGCCAGCTCGACTGGACCGAGAGATCCGGACACAGCGGCTTCTGTG 399
DB 337 GAACACAGCCTCAGCCAGCTCGACTGGACCGAGAGATCCGGACACAGCGGCTTCTGTG 396
QY 400 ACCATCCCGGTACTCTGGAGGCTCGACGAGCAGCAGCAGCAGCTCCGGCTCGCTGGG 459
DB 397 ACCATCCCGGTACTCTGGAGGCTCGACGAGCAGCAGCAGCAGCTCCGGCTCGCTGGG 456
QY 460 CTGGCGCTGCCCGTCC 475
DB 457 CTGGCGCTGCCCGTCC 472

RESULT 4
LOCUS AI394679/c
DEFINITION tg24b03.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2109677 3'
similar to SW:TRY3_CHICK Q90629 TRYPSINOMEN 2-P29 PRECURSOR.
; contains TARI.B1 TARI repetitive element ; mRNA sequence.
AI394679
VERSION AI394679.1 GI:4224226
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 469)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

QY 523 ACCGCTGGCAGCGAGTCCAGCTCTCAGGCTGGGGCATACCAACCAACCAACCAACCAACCA 582
DB 358 ACTACAGGGGCCATGATGCTCATGCTCAGATGGGTACTACAACCAACCAACCAACCAAC 417
QY 583 TTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCATGCTCCATGCTCCATGCT 642
DB 418 TTCCCGAGCGGCTCCAGTGCCTCAACCTCTCCATGCTCCCATGCTCCATGCTCCATGCT 477
QY 643 GTGTATCCCGGAGATCAGAGCAATGCTGTGTGTCAGCGCGCGCTCCCGGGCAGGAT 702
DB 478 GTGTATCCCGGAGATCAGAGCAATGCTGTGTGTCAGCGCGCGCTCCCGGGCAGGAT 537
QY 703 GCCTGCGCAGGATGATCTGGGGGCCCCCTGCTGTGTGGGGGAGTCTCTCAAGTCTGGTG 762
DB 538 GCCTGCGCAGGATGATCTGGGGGCCCCCTGCTGTGTGGGGGAGTCTCTCAAGTCTGGTG 597
QY 763 TCCTGGGGGCTGTGGGGGCCCCCTGCTGTGTGGGGGAGTCTCTCAAGTCTGGTG 822
DB 598 TCCTGGGGGCTGTGGGGGCCCCCTGCTGTGTGGGGGAGTCTCTCAAGTCTGGTG 657
QY 823 TGCAGATGCTGGAGTCCCGATGATCATGAGGAACTGACCTGTTCTCCACCC 882
DB 658 TGCAGATGCTGGAGTCCCGATGATCATGAGGAACTGACCTGTTCTCCACCC 717
QY 883 TCACACCCCGCTTAACTTGGGTACCCCTCTGCGCCCTCAGAGCAACATATCTCTCC 942
DB 718 TGCA---TGCGAGTCTGCTTGGGAGCGGCTCTGACCTCGAGCAACAAAGTCTCTCT 773
QY 943 ATGACTTCCCTAGCTCAGCTCTTGTGGCTGGGAGTCTCTGGAAT 991
DB 774 CTCACCACTCCAGCTCCCGCTTGTGTGAGCTTAGAACCTCTTGAACCT 822

RESULT 3
LOCUS BX110362
DEFINITION BX110362 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2109677 3'
IMAGE:2109677, mRNA sequence.
BX110362
VERSION BX110362.1 GI:27836332
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 472)
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
JOURNAL Human Unigeneset - RZPD3
COMMENT Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGp998E065203.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.

FEATURES
source
Location/Qualifiers
1..472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998E065203 ; IMAGE:2109677"

found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1051 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 397.

FEATURES

1. 469 Location/Qualifiers

1. 469 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2109677"
 /tissue_type="B-cell, chronic lymphocytic leukemia"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP CLL1"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTCGGAGCGCGCATGCTCTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 28.6%; Score 311.8; DB 9; Length 469;
 Best Local Similarity 97.3%; Pred. No. 96-50; 7; Indels 2; Gaps 1;
 Matches 328; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
 QY 493 CAACCCCTGCCCTGCCAATGACTGTGCAACCGTGGCAGCGAGTGCACGCTTCAGGC 552
 Db 469 CAACCCCTGCCCTGCCAATGACTGTGCAACCGTGGCAGCGAGTGCACGCTTCAGGC 410
 QY 553 TGGGCATCACCAACACCCAGAACCCATCCCGATCTCTCCAGTCCCTCAACCTC 612
 Db 409 TGGGCAT--CCCAACCAACCCCGAACCCATCCCGATCTCTCCAGTCCCTCAACCTC 352
 QY 613 TCCATGCTCCCATGCCACCTGCCATGCTGTATCCCGGAGATACAGAGCAACATG 672
 Db 351 TCCATGCTCCCATGCCACCTGCCATGCTGTATCCCGGAGATACAGAGCAACATG 292
 QY 673 GTGTGTGACGGCGGTCCCGGAGGAGTGCCTGCGAGGTGATTCGGGGCCCCCTG 732
 Db 291 GTGTGTGACGGCGGTCCCGGAGGAGTGCCTGCGAGGTGATTCGGGGCCCCCTG 232
 QY 733 GTGTGTGGGGAGTCTTCAAGTCTGTCTGCGGTCTGTGGGGCCCTGTGACAA 792
 Db 231 GTGTGTGGGGAGTCTTCAAGTCTGTCTGCGGTCTGTGGGGCCCTGTGACAA 172
 QY 793 GATGGCATCCCTGGAGTCTACACCTATATTGCAAGT 829
 Db 171 GATGGCATCCCTGGAGTCTACACCTATATTGCAACT 135

RESULT 5

BQ195790/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 676)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

97044477

PUBMED

COMMENT

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized fundus library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
 Seq primer: M13 Forward

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1. 676

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-CN1-cmq-p-20-0-UI"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="UI-R-CN1"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CN1 library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, and normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CN1 library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CAO and Cal corresponding to plates R-CAO-AWV through R-CAO-AXS, R-CAO-AZX through R-CAO-BAZ, R-CAO-BFE through R-CAO-BHY, R-CAO-BJS, R-CAO-BKE, R-CAO-BKG-H, R-CAO-BKJ-K, R-CAO-BKP through R-CAO-BKS, R-CAO-BKU-V, R-CAO-BLY through R-CAO-BMA, R-CAO-EMC through R-CAO-BME, R-CAO-BNS, R-CAO-BOB through R-CAO-BOJ, R-CAO-BPA through R-CAO-BPG, R-CAI-BJR, R-CAI-BDA, R-CAI-BHZ through R-CAI-BJF, R-CAI-BKJ, R-CAI-BKT, R-CAI-BKB, R-CAI-BKF, R-CAI-BKH, R-CAI-BKI, R-CAI-BKF, R-CAI-BLF, R-CAI-BLH through R-CAI-BLN, R-CAI-BLS, R-CAI-BLU-V, R-CAI-BNR, and R-CAI-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, and CX0s and normalized libraries CS0, CT0, CU0, CWO, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCE through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BID through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CWO-BVY through R-CWO-BWV, R-CWO-BXN through R-CWO-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates

R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer including CY0, CZ0, DAO, DB0, DCO, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYL, R-CZ0-BZB-C, R-DAO-BYJ through R-DAO-BYP, R-DAO-BZD through R-DAO-BZH, R-DB0-BYQ through R-DB0-BZA, R-DCO-BZ1 through R-DCO-BZQ, R-DCO-CAY through R-DCO-CBA, R-DD0-BZR through R-DD0-CA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLM-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CNO pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkx-a-05-0-UI, bkx-a-06-0-UI, bkx-a-11-0-UI, bkx-c-06-0-UI, bkx-c-09-0-UI, bkx-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-05-0-UI, blb-f-02-0-UI, blb-f-02-0-UI, blg-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CNO clone (corresponding to the address bkx-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_TISSUE=fundus
TAG_LIB=UI-R-CN1
TAG_SEQ=TTCCG

ORIGIN

Query Match 28.0%; Score 305.2; DB 13; Length 676;
Best Local Similarity 67.6%; Pred. No. 1.6e-48;
Matches 469; Conservative 0; Mismatches 203; Indels 22; Gaps 2;

QY 398 TGACCCATCCCGGTACTCGAGGCTCGAGGCGCAGAGCAGACCTCGGCTGCTGC 457
DB 674 TAAACACCCCGGTACTCGAGGCTCGAGGCGCAGAGCAGACCTCGGCTGCTGC 615

QY 458 GCGTCGCTGCGCTGCGGTAAACAGCAGGCTTCAACCCCTGCGCTGCGGCTGCTG 517
DB 614 GCGTGAACAGACCCATAGCTTGAATGCTGCTGCGGCGCTGCGCTGCGGCTGCTG 555

QY 518 GTGCACCCGCTGCGCAGGCTGCGGCTGCTGAGGCTGCGGCTGCGGCTGCGGCTG 577
DB 554 GTGCACCCGCTGCGGCTGCGGCTGCTGAGGCTGCGGCTGCGGCTGCGGCTG 495

QY 578 ACCCATTCGCGATGCTTCCAGTGGCTCAACTCTCCATGCTTCCATGCGGCTGCG 637
DB 494 ACCCATTCGCGATGCTTCCAGTGGCTCAACTCTCCATGCTTCCATGCGGCTGCG 435

QY 638 ATGGTGCTATCCCGGGAATCAGAGCAACATGCTGCTGCTGCTGCTGCTGCTG 697
DB 434 GGGCCGATTTCTTGAAGAGTGACAGAGAAATGCTGCTGCTGCTGCTGCTGCTG 375

QY 698 AGGATGCTGCCAGGCTGATCTTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 757

DB 374 AGATGCTTGTGAGGTGACTCTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG 315
QY 758 TGGTGTCTCTGGGGGTCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 817
DB 314 TGGTGTCTCTGGGGGTCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 255

QY 818 ATATTTGCAATGATGAGGCTGAGTCCGATGATGATGATGATGATGATGATGAT 877
DB 254 AGGTCTGCAATACACGGACTGGATCAGATGATGATGATGATGATGATGATGAT 195
QY 878 CCACCTCCACCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 937
DB 194 CGGTCTGATGAGCTCTGCTTA-----GGGCACTCCGACCTCGGAGCACTGCT 140
QY 938 CTCTCCATGCTTCCCTAGTCTTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCT 997
DB 139 GTCCCTCTGGCACTCCCAAGCTCCCACTCTTGGAGCTTAG-----AACC 97

QY 998 CTGCGGAGGCTTCTTAAGAGCCAGGCGGGGGGGGGGGGGGGGGGGGGGG 1057
DB 96 TCTGTAAACCTTGGTGTATCTATGTCAGTGGAGGGGGGGGGGGGGGGGGGG 37

QY 1058 AATATAAATGAAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1091
DB 36 AAGTTATACACAGGGGAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3

RESULT 6
BX514743 520 bp mRNA linear EST 25-JUN-2003
LOCUS
DEFINITION BX514743 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGP998N221098; IMAGE:464085, mRNA sequence.
ACCESSION BX514743
VERSION BX514743
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Heil.O., Ebert.L., Neubert.P., Peters.M., Radelof.U., Schneider.D.
and Korn.B.
Mouse UnigeneSet - RZPD2
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGP998N221098.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)
http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl/cgi/response?libNo=981 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
T7, Primer sequence: TAATACGACTCACTATAGG.

FEATURES
source
1. 520
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGP998N221098 ; IMAGE:464085"
/dev stage="19.5 dpc total fetus"
/lab host="DH10B (ampicillin resistant)"
/clone.lib="Soares mouse p3NMF19.5"
/note="Vector: pMT73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'

TGTTACCACTCTGAAGTGGAGCGCGGCATTTTCTTTTCTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Patima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 25.9%; Score 282.6; DB 13; Length 520;
Best Local Similarity 72.7%; Pred. No. 3.6e-44;
Matches 380; Conservative 0; Mismatches 139; Indels 4; Gaps 1;
QY 448 CGGTGTCTGGCGCTGCGCGCTGCGCGTACCAAGACGGTTCACCCCTGCCCGTG 507
Db 2 CGACTCTCGCGCTGAACAGACCCCATCCACCTGACCGCTGCTGCTCGCGCGCTGCGCGCTG 61
QY 508 CCCAATGACTGTGCAACCGCTGGCAGCGAGTGCACGCTCTCAGGTGGGCGATCACCAAC 567
Db 62 CCCAGTTCCTGTGTAACACAGGGGCGCATGTCTCATGTCTCAGATGGGGTACTACAAAC 121
QY 568 CACCACAGGAACCCATTCCTCCGATCTGTCTCAGTGCCTCAACCTCTCCATCTCTCCAT 627
Db 122 AAGCCATGGGACCATTCACAGACGGCTCCAGTGCCTCAACCTCTCCATCTCTCCAT 181
QY 628 GGCACCTGCTGATGTGTATCCGGGAGATACAGGACACATGTGTGTGCGAGGCGC 687
Db 182 GAGACGTGCGCGCTGTGTTTCTCGGAAGAGTACGAGAAATATGTGTGCGAGGTGA 241
QY 688 GTCCCGGGGAGGATGCTGCGAGGTCATCTGCGGGGCCCTGCTGTGCGGGGAGTC 747
Db 242 GAACCGGGAAGGATGCTGTGAGGTGACTCTGAGAGGCCCTGCTGTGTGGAGGGTT 301
QY 748 CTTCAAGGTCTGTGCTCTGGGGGCTGTGTGGGGCCCTGTGGAACAGATGGCATCCCTGA 807
Db 302 CTTCAAGGTCTGTGCTCTGGGGTCTGTGGGGCTGTGTGTCATGATGATTCAGGA 361
QY 808 GTCTACACTATATTTGCAATGTGACTGTGATCCGATCGGATGATCATGAGCAACTGA 867
Db 362 GTCTACAGAGGTGTGAAAATACACGAGCTGGATCAGAAATAGTATCAGGAATAACTAA 421
QY 868 CTTGTTCTCTCCACTCCACCCACCCCTTAACTTGGGTACCCCTCTGCGCCCTCAGAGC 927
Db 422 GTGACTCCTCAGCTGCA----TGCAGCTCTGTGTAGGACCGCTCTGACCTTCGAGC 477
QY 928 ACCAATATCTCTCCATCACTTCCCTAGTCCACTCTGTGTG 970
Db 478 ACCAAGTCTCTCTCCTCCTCACCCTCAGGCTCCACCTTGTG 520

RESULT 7
AI324852/c
LOCUS mi20d11.x1 Soares mouse p3NM19.5 Mus musculus cDNA clone
DEFINITION IMAGE:464085 3' similar to SW:TRY1_CHICK Q90627 TRYPSINOGEN P1
PRECUSOR. ; mRNA sequence.

ACCESSION AI324852
VERSION AI324852.1 GI:4059281
KEYWORDS EST.

SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 545)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilton, R. and
Waterston, R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:277901

This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
High quality sequence stop: 475.

FEATURES
source

1..545
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:464085"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NM19.5"
/notes="vector: p773D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Site 3: Eco RI;
was primed with a Not I - oligo(dT) primer [5',
TGTTACCACTCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Patima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 24.7%; Score 269.4; DB 9; Length 545;
Best Local Similarity 72.3%; Pred. No. 1.2e-41;
Matches 365; Conservative 0; Mismatches 136; Indels 4; Gaps 1;
QY 516 CTGTGCAACCGCTGGCAGCGCTCCAGTCCAGTCCCACTCTCCATCGTCTCCATGCCACCTG 575
Db 541 CTGTGTAACATACAGGGCGCATGTGTCTATGTCTCAGGATGGGGTACTACAAACAGCCATG 482
QY 576 GAAACCCATTCCTCCGATCTGTCTCCAGTGCCTCAACCTCTCCATCGTCTCCATGCCACCTG 635
Db 481 GGACCCATTCCTCCAGCGGCTCCAGTGCCTCAACCTCTCCATCGTCTCCATGCCACCTG 422
QY 636 CCATGTGTGTATCCCGGAGAAATACAGAGCAATGTGTGTGTCAGGCGGCTCCCGG 695
Db 421 CCGGCTGTGTCTCTGGAAGAGTGACGGAGATATGTGTGTGAGGTGGAGAGCGCG 362
QY 696 GCAGATGCTCCAGGGGTGATTCCTGGGGGCCCTGGTGTGTGGGGAGTCTCTTCAAGG 755
Db 361 GAAGATGCTCCAGGGGTGATTCCTGGGGGCCCTGGTGTGTGGGGAGTCTCTTCAAGG 302
QY 756 TCTGTGTCTCCGCGGTCTGTGGGGCCCTGTGGAACAGATGCCATCCCTGAGGTCTACAC 815
Db 301 TCTGTGTCTCCGCGGTCTGTGGGGCCCTGTGGAACAGATGCCATCCCTGAGGTCTACAC 242
QY 816 CTATATTGCAAGTATGTGGAGTCCGATCGGATGATCATGAGGAACAATGACCTGTTC 875
Db 241 GAAGGTCTGCAAAATACACGGAGTGGATCAGAAATAGTATCAGGAATACTAAGTGACTCC 182
QY 876 TCCACCTCCACCCCGCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCAACCAATAT 935
Db 181 CTCAGCTGTATGAG----CTCTGCTTAGGAGCGCTCTGACCTCTCGAGGACCAAGGT 126
QY 936 CTCCTCCATCACTTCCCTAGCTCCACTCTTGTGTGGCTGGGAACTCTTGGAACTTTAA 995
Db 125 CTCCTCCCTCACCCTCCAGGCTCCACCTTGTGTGAGCTTAGAACTCTTGAACCTCGA 66
QY 996 CTCCTGCCAGCCCTTCTTAAGCCCCA 1020
Db 65 CTCACCCATGGCAGCGGAGGCA 41

QY 628 GCCACCTGCCATGCTGTATCCCGGAGATCAGAGCAACATGGTGTGTGCGAGCGGC 687
 DB 180 GAGACGTGCGGCGTGTGTTCCTGGAGAGTGCAGGAGAAATATGTTGTGTGCGAGTGA 239
 QY 688 GTCCCGGCGGAGAGTGCCTGCCAGGAGTATCTGGGGGCCCCCTGGTGTGTGGGGAGTC 747
 DB 240 GAAGCGGAGAGAGTGCCTGTGACGGTGACTCTGGAGGCCCCCTGGTGTGTGGAGGGTT 299
 QY 748 CTTCAAGGTCTGGTGTCTGTGGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCCTGA 807
 DB 300 CTTCAAGGTCTGGTGTCTGTGGGGTCTGTGGGGCCCTGTGGTCAAAAAGGTATCCAGA 359
 QY 808 GTCTACACTATATTTGCAAGTATGTGACTGGATCCCGATGATCATCAGGAACAACATGA 867
 DB 360 GTCTACAGAGAGTCTGCAATACAGGACTGGATCAGATAGTATCATCAGGATTAATA 419
 QY 868 CTTGTTCTCTCACTCCACCCCAACCCCTTAACCTGGTGGTACCCCTGTGGCCCTCAGAGC 927
 DB 420 GTGACTCCCTCAGCTGCA---TGCAGCTCTGCTTAGGGACCGCTCTGACCTCGGAGC 475
 QY 928 ACCATATCTCTCTCCATCATCTCCCTAGCT 958
 DB 476 ACCAAGGTCTCTCTCTCCTCACCACCTCAGGCT 506

RESULT 9
 AI322408 449 bp mRNA linear EST 23-DEC-1998
 LOCUS mi20d11.y1 Soares mouse p3NNF19.5 Mus musculus cDNA clone
 DEFINITION IMAGE:464085 5' similar to SW:TRY3_SALSA P35033 TRYPSINOGEN III
 PRECURSOR ;, mRNA sequence.

ACCESSION AI322408
 VERSION AI322408.1 GI:4056837
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 449)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:277901
 This read is a RESSEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Putative full length read
 vector to vector length is 607
 Seq primer: -40RP from Gibco
 High quality sequence stop: 408.

FEATURES
 source
 1..449
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:464085"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NNF19.5"
 /note="vector: p1713D (Pharmacia) with a modified

RESULT 8
 AA028356 506 bp mRNA linear EST 11-SEP-1996
 LOCUS mi20d11.r1 Soares mouse p3NNF19.5 Mus musculus cDNA clone
 DEFINITION IMAGE:464085 5' similar to SW:TRY3_BOVIN P00760 TRYPSINOGEN ;, mRNA
 sequence.

ACCESSION AA028356
 VERSION AA028356.1 GI:1494434
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 506)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:277901
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 486.

FEATURES
 source
 1..506
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:464085"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NNF19.5"
 /note="vector: p1713D (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p1713 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Patima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."

ORIGIN
 Query Match 24.6%; Score 268.2; DB 9; Length 506;
 Best Local Similarity 73.0%; Pred. No. 2.1e-41;
 Matches 373; Conservative 0; Mismatches 133; Indels 5; Gaps 2;
 QY 448 CGGCTGTCGGCTGGCTGCCCTCGCGGTAACAGCAGGTTCAACCCCTGCCCTG 507
 DB 1 CCACTCTCGCGCTGAAACAGACCCATCACTGACCGCTGTCGCGCCGCTG-GCCTG 59
 QY 508 CCCAATGATGTGCAACCGCTGGCACCGAGTCCACGTTCTCAGGCTGGGGCATCCCAAC 567
 DB 60 CCCAGTCTCTGTGTAACACAGGGCCATGTGTGATGCTCAGGATGGGTACTACAAAC 119
 QY 568 CACCACGGAACCCATTCCCGGATCTGCTCCAGTCCGCTCAACTCTCCATGCTCCCAT 627
 DB 120 AAGCCATGGGACCCATTCCACAGCGGCTCCAGTGCCTCAACCTCTCCACTGCTCCAAT 179

RESULT 10	BC035385	1186 bp	linear	HTC 04-MAR-2003
LOCUS	BC035385			
DEFINITION	Homo sapiens, Similar to kallikrein 12, clone IMAGE:4825490, mRNA.			
ACCESSION	BC035385			
VERSION	BC035385.1	GI:23958876		
KEYWORDS	HTC			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1186)			
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian			
	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
	USA			
REMARK	NTH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk			
	Email: cgapbs-remail.nih.gov			
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
	cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki			
	Toshiyuki and Piero Carninci (RIKEN)			

COMMENT

COMMENT

Db 640 CCAGTGCCTCATTTCTGGATGGGGCAACCAAGTCCAGCCCCCAGTTGGCGCTGCCTCAT 699
QY 593 TGTCCAGTGCCTCAACCTCTCCATGGTCTCCCATGCCACCTGCCATGGTGTGTATCCCG 652
Db 700 CCTTGCATGTGCCAATGTCTCCATCATCGAACACAGGAGTGTGAGAGGCCCTACCCCG 759
QY 653 GGAGAAATCAGAGCAACATGGTGTGTGAGGCG---GCGTCCCAGGCGAGGATGCCCTGCC 709
Db 760 GCAACATCACAGACACCATGCTGTGGCCAGTGTTCGAAAGAGGGCAAGGACTCCTGTC 819
QY 710 AGGGTGATTCCTGGGGCCCCCTGGTGTGGGGAGTCTTCAAGTCTGGTGTCTCTGGG 769
Db 820 AGGGTGAATCTGGAGCCCCCTGGTGTGCAACGGATCTCTTCAAGGATCATCTCTCTGGG 879
QY 770 GGTCTGTGGGGCCCTGTGGAACAAGATGGCATCCCTGGAGTCTACCTATATTGCAAGT 829
Db 880 G---TCAGGACCCCATGTGCGGTCAACAGAAAGCCTGGTGTCTATACAAAAGTCTGCAAT 936
QY 830 ATGTGACTGGATCCGGATGATCATGAGGAACAACTGACCTGTTCTCCACCTCCACCC 889
Db 937 ACTTTAACTGGATCCACGAGGTTATGAGGAACAATTAGAGGGGACCTGCTTCCACCCACC 996
QY 890 CCACCCCT 897
Db 997 CAACCCCT 1004

Search completed: June 20, 2004, 04:42:18
Job time : 3358 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 02:20:26 ; Search time 99 Seconds
(without alignments)
6115.671 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	506.4	46.4	840	4	US-09-244-111-5
2	493.2	45.2	619	4	US-09-280-116-139
3	219.4	20.1	1146	4	US-09-205-258-247
4	217.8	20.0	1166	3	US-08-944-483-7
5	217.6	19.9	1192	3	US-08-944-483-8
6	217.6	19.9	1314	3	US-09-025-059-2
7	216.8	19.9	1292	4	US-09-205-258-189
8	206.8	19.0	833	2	US-08-730-137-2
9	199.4	18.3	1052	4	US-09-386-642-10
10	191.6	17.6	1343	4	US-09-618-259-72
11	191.6	17.6	1360	4	US-09-618-259-6
12	190	17.4	994	3	US-09-008-271A-19
13	188.4	17.3	1049	4	US-09-386-642-9
14	184.8	16.9	944	3	US-09-070-526-1
15	173.8	15.9	1454	1	US-08-467-155A-2
16	173.8	15.9	1454	2	US-08-628-198-2
17	173.8	15.9	1454	3	US-09-201-038-2
18	173.8	15.9	1454	5	PCT-US96-07343-2
19	172.8	15.8	825	3	US-09-120-582-1
20	169.4	15.5	897	2	US-08-956-267A-1
21	162	14.8	732	1	US-08-361-395-2
22	146	13.4	1499	4	US-09-509-908-1
23	143	13.1	969	3	US-09-502-600-30
24	143	13.1	969	4	US-09-918-243-30
25	143	13.1	986	2	US-08-557-146-1
26	143	13.1	986	2	US-08-154-344-1
27	143	13.1	1089	3	US-08-930-188-1

c	28	143	13.1	1089	3	US-08-930-188-3	Sequence 3, Appli
c	29	143	13.1	1089	5	PCT-US96-04294-1	Sequence 1, Appli
c	30	143	13.1	1089	5	PCT-US96-04294-3	Sequence 3, Appli
	31	136	12.5	693	4	US-09-402-515A-17	Sequence 17, Appli
	32	135	12.4	1476	2	US-08-824-874-2	Sequence 2, Appli
	33	135	12.4	1476	3	US-09-210-084-2	Sequence 2, Appli
	34	135	12.4	1476	4	US-09-764-762-2	Sequence 2, Appli
	35	125.8	11.5	725	3	US-09-197-801-12	Sequence 12, Appli
	36	125.8	11.5	725	3	US-09-551-028-12	Sequence 12, Appli
	37	125.8	11.5	725	4	US-09-664-595A-12	Sequence 12, Appli
	38	122.2	11.2	792	4	US-09-244-111-3	Sequence 3, Appli
	39	118.6	10.9	1504	4	US-09-280-116-1	Sequence 1, Appli
	40	118	10.8	1341	4	US-08-983-075D-6	Sequence 6, Appli
	41	118	10.8	1358	4	US-08-983-075D-8	Sequence 8, Appli
	42	117.4	10.8	832	3	US-08-768-859A-5	Sequence 5, Appli
	43	117.4	10.8	832	3	US-08-768-859A-20	Sequence 20, Appli
	44	117.4	10.8	832	3	US-08-767-820A-5	Sequence 5, Appli
	45	117.4	10.8	832	3	US-08-767-820A-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1

US-09-244-111-5
; Sequence 5, Application US/09244111
; Patent No. 6566498
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides
; FILE REFERENCE: PF391
; CURRENT APPLICATION NUMBER: US/09/244,111
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 60/073,961
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)..(603)
US-09-244-111-5

Query Match 46.4%; Score 506.4; DB 4; Length 840;						
Best Local Similarity 97.9%; Pred. No. 3.4e-130;						
Matches 513; Conservative 0; Mismatches 11; Indels 0; Gaps 0;						
Qy	568	CACCCACGGAAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCAT	627			
Db	304	CAGCGGACAGCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCAT	363			
Qy	628	GCCACCTGCCATGGTGTGTATCCGGGAGATCAACAGCAACATGTGTGTGACGGCGGC	687			
Db	364	GCCACCTGCCATGGTGTGTATCCGGGAGATCAACAGCAACATGTGTGTGACGGCGGC	423			
Qy	688	GTCCCGGGGACAGATGCTCCAGGTGATCTGGGGGCCCCCTGGTGTGGGGGAGTC	747			
Db	424	GTCCCGGGGACAGATGCTCCAGGTGATCTGGGGGCCCCCTGGTGTGGGGGAGTC	483			
Qy	748	CTTCAAGGTCTGGTGTCTCTGGGGGTCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGA	807			
Db	484	CTTCAAGGTCTGGTGTCTCTGGGGGTCTGTGGGGCCCTGTGGCAAGATGGCATCCCTGGA	543			
Qy	808	GTCTACACCTATATTTGCAAGTATGTGAGATCGGATCGGATGATCATGAGGAACAATGA	867			
Db	544	GTCTACACCTATATTTGCAAGTATGTGAGATCGGATCGGATGATCATGAGGAACAATGA	603			
Qy	868	CTGTGTTCTTCCACCTCCACCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGC	927			
Db	604	CTGTGTTCTTCCACCTCCACCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGAGC	663			

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 247
LENGTH: 1146
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (20)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (35)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (36)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (37)
OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-247

Query Match 20.1%; Score 219.4; DB 4; Length 1146;
Best Local Similarity 58.5%; Fred. No. 7.1e-51;
Matches 436; Conservative 3; Mismatches 294; Indels 12; Gaps 3;
QY 168 GGCAGGCACACCGAAGATTTTCAATGGCACTGAGTGGGGCGTAACCTACAGCCGGTGGCA 227
DB 108 GGGGAGAGACCCAGGATCATCAGGGTTGAGTGCAGCCCTCACTCCAGCCCTGGCA 167
QY 228 GTGGGGCTTTTGGGGCAACAGCCTGCGCTCGGGGGTGTCTTATTGACACACAGGTG 287
DB 168 GGCAGCCCTTTCGAGAAGACGGGGTACTCTGTGGGGCGAGCTCATCGCCCCAGATG 227
QY 288 GGTCTTCACAGCGGCTCACTGCGGCGCAGAGGACTGGGTGGCTGGGGGAACACAG 347
DB 228 GCTCTTCACAGCGCCCACTGCTCAAGCCCGCTACATAGTTTCACTCGGGGGCAGACAA 287
QY 348 CTTGAGCAGCTCGATGGACCGAGCAGATCCGCGGCTTCTCTGTGACCCATCC 407
DB 288 CTTCCAGAGGAGGGCTGTGAGCAGACCCGAGCAGCCAGTGTCTTCCCCACCC 347
QY 408 CGGCTACCTGGGAGCCTTCGACGAGCCACGAGAC-----GACTCGGCTGTGGGCT 461
DB 348 CGGCTTCAACACAGCCTTCCCAACAAAGACCCGCAATGACATCATGTCTGTGAAGAT 407
QY 462 GCGCTGCGGCTCGGCTATACAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGC 521

DB 408 GGCATCGCCAGTCTCCATCACCTGGGCTGTGGACCCCTCACCCCTCTCCTCAGCTGTGT 467
QY 522 AACCGTGCACGAGTGCACGTCTCAGGCTGGGGCATCAACACACCCACCGAACCC 581
DB 468 CACTGTGTGCACCGAGCTGYCTCATTTCCGGCTGGGGCAGMACGTCCAGGCCCGGTTACG 527
QY 582 ATTCCCGGATCTGCTCAGTGCCTCAACCTCTCCATCGTCTCCCATGCCACCTGCCATGG 641
DB 528 CTTGCTCACACTTGGGATGCCCAACATCACCATCATTTGACACACAGAGTGTGAGAA 587
QY 642 TGTGTATCCCGGAGAAATCAGAGCAACATGTGTGTGAGGGG---GGTCCCGGGGCA 698
DB 588 CGCCTACCCCGGCAACATCACAGACACCATGTGTGTGCGCAGCGTGCAGGAAGGGGCA 647
QY 699 GATGCTGCCAGGATGATCTGG 758
DB 648 GGACTCTCCAGGGTGAATCTGG 707
QY 759 GGTGTCTGGGGTCTGTGG 818
DB 708 TATCTCTGGG---CCAGATCCGTGTGGGATCACCCGAAAGCCCTGGTGTCTACAGAA 764
QY 819 TATTGCAAGTATGTGAGCTGGATCCGATGATCATGAGGAAACATGACCTGTTCCTC 878
DB 765 AGTCTGCAATATGTGAGTGGATCCAGGAGACGATGAAGAAACATTAGACTGGACCCAC 824
QY 879 CACTCCACCCCGGCTTAACTT 903
DB 825 CCACACAGCCCATCACCTCCATT 849

RESULT 4

US-08-944-483-7
Sequence 7, Application US/08944483
Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729

; TELEFAX: 847/938-2623

; TELEX:

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1166 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-944-483-7

Query Match 20.0%; Score 217.8; DB 3; Length 1166;

Best Local Similarity 58.9%; Pred. No. 2e-50;

Matches 432; Conservative 1; Mismatches 289; Indels 12; Gaps 3;

```
QY 179 CGAAGATTTCAATGGCACTAGTGTGGGGTAACTCAACGGTGGAGTGGGGCTGT 238
Db 161 CCAGGATCATCAAGGGTTTCAGTGCNAGCCCTCACTCCAGCCCTGGCAGGAGCCCTGT 220
QY 239 TTGAGGGACACGAGCTCGCTGGGGGTGCTCTTATTGACACAGGTGGTCTCTCAG 298
Db 221 TCRAGAARACGGGTATCTGTGGGGGAGCGCTCATCGCCCGAGATGGCTCTCTGACG 280
QY 299 CGGCTCACTGACGGCGAGCAAGTACTGGTGGCGCTGGGGGAAACACAGCTTCAGCCAGC 358
Db 281 CAGCCCACTGCTCAAGCCCGCTACATAGTTCACTGGGGCAGCACAACTCCAGAAAG 340
QY 359 TCGACTGACGAGCAGATCGGACACAGGGCTTCTGTGACCCATCCCGGCTACCTGG 418
Db 341 AGGAGGGTGTGAGCAGACCGGACAGCCATGAGTCTCTCCCGACCCCGGCTTCAACA 400
QY 419 GAGCTCCAGCAGGACACAGCAGC-----GACCTCCGGCTGTGGGGTGGCGCTGCCG 472
Db 401 ACAGCTCCCCCAACAAAGACACCGCAATGATCATCTGTGTAAGATGGCATCGCCAG 460
QY 473 TCCGGGTAAACAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCGCTGGCA 532
Db 461 TCTCCATCACCTGGGCTGTGGGACCCCTCACCTCTCTCTCAGCTGTGTCTACTGTGGCA 520
QY 533 CCGAGTGCACGTCTCAGCTGGGGCATCAACACCCAGCAGCAACCCATTCCCGGATC 592
Db 521 CAGCTGCTCATTTCCGGCTGGGGCAGCAGCTCCAGCCCGCAAGTACGCTGCTCTACA 580
QY 593 TGCTCCAGTGCTCAACCTCTCCATCGTCTCCATGCGCAGCTGCCATGGTGTGTATCCG 652
Db 581 CTTTGGATGCGCCAAACATCACTATTGAGCACCAGAGTGTGAGAAACGCTTACCOCG 640
QY 653 GGAGNATCAGAGCAACATGTTGTGTGAGGG-----GGTCCCGGGGAGGATGCTGCC 709
Db 641 GCAACATCAGACACCAATGTTGTGTGCGAGCGTGCAGGAAGGGGGCAAGACTCTGCC 700
QY 710 AGGGTATTTCTGGGGGCCCCCTGGTGTGTGGGGAGTCTCTCAAGGTCTGTGTCTGGG 769
Db 701 AGGGTATCTCGGGGGCCCTCTGGTCTGTAAACAGTCTCTTCAAGGCATATCTCTGG 760
QY 770 GGTCTGTGGGGCCCTGTGGAAGAAGTGGATCCCTGGAGTCTACACCTATATTGCAAGT 829
Db 761 G---CCAGGATCCGCTGTGCGATCACCCGAAAGCCCTGGTGTCTACACGAAAGTCTGCAAT 817
QY 830 ATGTGAGTGGATCCGATGATCATGAGGAAACACTGACCTGTTCTCTCCACCTCCACCC 889
Db 818 ATGTGAGTGGATCCGAGGACGATGAAGAAACAATTAGACTGGACCCCAACCAACCAAGCC 877
QY 890 CCACCCCTTAATT 903
Db 878 CATCACCTCCATT 891
```

RESULT 5

US-08-944-483-8

; Sequence 8, Application US/08944483

; Patent No. 6232456

; GENERAL INFORMATION:

; APPLICANT: COHEN, MAURICE

```
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183-US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-944-483-8
```

Query Match 19.9%; Score 217.6; DB 3; Length 1192;

Best Local Similarity 59.0%; Pred. No. 2.3e-50;

Matches 433; Conservative 0; Mismatches 289; Indels 12; Gaps 3;

```
QY 179 CGAAGATTTCAATGGCACTAGTGTGGGGTAACTCAACGGTGGAGTGGGGCTGT 238
Db 165 CCAGGATCATCAAGGGTTCAGTGCNAGCCCTCACTCCAGCCCTGGCAGGAGCCCTGT 224
QY 239 TTGAGGGACACGAGCTCGCTGGGGGTGCTCTTATTGACACAGGTGGTCTCTCAG 298
Db 225 TCGAGAAGACGGGCTACTCTGTGGGGGAGGCTCATCGCCCGCAGATGGCTCTGACAG 284
QY 299 CGGCTCACTGACGGCGAGCAGGTACTGGGTGGCGCTGGGGGAAACACAGCTTCAGCCAGC 358
Db 285 CAGCCCACTGCTCAAGCCCGCTACATAGTTTACCTGGGGGAGCAAACTCCAGAAGG 344
QY 359 TCGACTGACCGAGCAGATCGGACACAGCGCTTCTGTGACCCATCCCGGCTACCTGG 418
Db 345 AGGAGGGCTGTGAGCAGACCCCGGACAGCCACTGAGTCTTCCCCCAACCCCGGCTTCAACA 404
QY 419 GAGCTTCGACGAGCCACGAGCAC-----GACCTTCGGGCTGTCTGGGCTGGCGCTGCCG 472
Db 405 ACAGCTCCCCCAACAAAGACACCGCAATGACATCATGCTGTGTAAGATGGCATCGCCAG 464
QY 473 TCCGGTAAACAGCAGGCTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCCCTGGCA 532
```


; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 189
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-205-258-189

Query Match 19.9%; Score 216.8; DB 4; Length 1292;
Best Local Similarity 58.7%; Pred. No. 3.9e-50;
Matches 431; Conservative 2; Mismatches 289; Indels 12; Gaps 3;

QY 179 CGAAGATTTCATATGGCACTGAGTGTGGGGGTAACTCAGAGCGTGGCAGGTGGGCTGT 238
DB 265 CCAGGATCATCAGGGGTTCGAGTCAAGCTCCTCAGCCCTGGCAGGCGCCTGT 324
QY 239 TTGAGGGCACCAAGCTTCGCTGCGGGGTGTCTTATTGACCACAGGTGGGTCTCACAG 298
DB 325 TCGAGAAGACGGGGTACTCTGTGGGGCGAGCTCATCGCCCCACAGATGGCTCTGACAG 384
QY 299 CGGCTCACTGACGGCAGCAGGTACTGGGTGGCTTGGGGGAAACAACAGCTCAGCCAGC 358
DB 385 CAGCCCACTGCCTCAAGCCCCGCTACATAGTTTCACTGGGGCAGCACAACTCCAGAAAG 444
QY 359 TCGACTGGACCGAGCAGATCCGGCAGCAGGGCTTCTGTGACCCCATCCCGGTCTCTGG 418
DB 445 AGAGGGGTGTGAGCAGACCCGAGCAGCCACTGAGTCTTCCCCCAGCCCGGTCTCAACA 504
QY 419 GAGCCTCGAGAGCCACGAGCAC- - - - -GACCTCGGGTGTGTGGGCTGCGGCTGCGCCG 472
DB 505 ACAGCCTCCCAACAAGACCCAGCAATGACATCATGTGTGTGAAGATGGCATCGCCAG 564
QY 473 TCGGGTAACACAGCAGGTTCAACCCCTGCCCCCTGGCCATGACTGTGCAACCGCTGGCA 532
DB 565 TCTCCATCACTGGGGTGTGCGACCCCTCACCTCTCTCCTCAGCTGTGTCACTGTGGCA 624
QY 533 CCGAGTGCCACAGTCTCAGGGTGGGGCATCACCAACACCCAGCAGCTCAGCCCTTCCCGGATC 592
DB 625 CCAGCTGCTCATTTCCGGCTGGGGCAGCAGCTCAGCCCTTACGCTGCTGCTCCTCACA 684
QY 593 TGCTCCAGTGCCTCAACCTTCCATCGTCTCCCATGCACTGCACTGCGATGGTGTGTATCCCG 652
DB 685 CTTTGGGATGCGGCAACATCACCATATTGAGCAGCAGAGTGTGAGAACGCTTACCCCG 744
QY 653 GGAGAAATCAGAGCAGCATGCTGTGCGAGG- - -CGTCCCGGGGCGAGGATGCGCTGCC 709
DB 745 GCAACATCAGACAGCACCATTGTTGTCAGCTGTGCGAGGAGGGGGGCAAGATCTCTGCC 804
QY 710 AGGTTGATTCTGGGGGCCCCCTGGTGTGGGGAGTCTTTCAAGTCTGGTGTCTGGG 769
DB 805 AGGTTGACTCGGGGGCCCTCTGTTGTGTAACCACTGCTTCAAGGCAATATCTCTGGG 864
QY 770 GGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTGGAGTCTACCTATATTTGCAAGT 829
DB 965 G- - -CCAGGATCCGTGTGCGATCACCCGAAAGCTGTGTGTCTACGAAAGTCTGCAAT 921
QY 830 ATGTGGACTCGATCCCGATCATGATGAGGAACAATGACCTGTTTCTCCACCTCCACCC 889
DB 922 ATGTGGACTGATCCAGGAGACCATGAGAACAAATTAGACTGGACCCACCAACAGGC 981
QY 890 CCACCCCTTAACCTT 903
DB 982 CATCACCTCCATT 995

RESULT 8

US-08-790-137-2
; Sequence 2, Application US/08790137
; Patent No. 5840871
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: A NOVEL PROSTATE-ASSOCIATED
; TITLE OF INVENTION: KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/790,137
; FILING DATE: Filed Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0195 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-790-137-2

Query Match 19.0%; Score 206.8; DB 2; Length 833;
Best Local Similarity 56.6%; Pred. No. 1.8e-47;
Matches 442; Conservative 0; Mismatches 327; Indels 12; Gaps 3;
QY 120 CATGGGGCTCAGATCTTTTGGCTCTCTGTGTCTTGGGCTCAGCCAGCCAGCCACACC 179
DB 35 CTGTGTTCTGTGCTCGCCCTGTCCCTGGGGGGGACTGGTGTGCGCCCGGATTCAATC 94
QY 180 GAAGATTTTCAATGGCACTGAGTGTGGGGGTAACTCAGAGCGGTGAGAGTGGGGCTGT 239
DB 95 CCGATTGTGGAGGCTGGAGTGTGAGAGCATTCAGAGCCCTGAGAGCGGCTCTGTA 154
QY 240 TGAGGGCACCAGCTCGCTGCGGGGGGTGTCTTATTGACCACAGAGTGGGTCTCTCAGC 299
DB 155 CCAGAAGACGCGCTACTCTGTGGGCGCAGCAGTCAATGCGCCCGGAGATGTTCTGACAGC 214
QY 300 GGCTCTACTGACGGGAGAGAGTACTGGGTGCGCTGGGGGAGAACAGAGCTCAGCCAGCT 359
DB 215 AGCCCACTGCTTNAAGCCCGCTACATAGTTTCACTGGGGGAGCAGCAACCTCAGAAGGA 274
QY 360 CGACTGAGCCAGCAGATCCGGCAGAGCGGCTTCTGTGAGCCATCCCGGCTACCTGGG 419
DB 275 GGAGGGCTGTGAGCAGACCCGAGCAGCAGCTAGTCTTCCCCCAGCCCGGCTTCAACAA 334
QY 420 AGCCTCGACAGCCACAGCAGC-----GACCTCCGGCTGCTCGGCTGCGCTGCGCGT 473
DB 335 CAGCCTCCCCAACAAAGACACCGCAATGACATCATGCTGGTGAAGATGGCATCGCCAGT 394
QY 474 CCGGTAAACAGCAGGTTTCAACCCCTGCCCTGCCCAATGATGTGCAACCGCTGGCAC 533

DB 395 CTCCATCACCTGGGCTGTGCGACCCCTCACCTCTCTCTCAAGCTGTGTCACTGTGGCAC 454
QY 534 CGAGTGCCACGCTCTCAGGCTGGGGCATCACCAACCCAGGAAACCCATTCGCGATCT 593
DB 455 CAGTGCCTCATTTCCGGCTGGGGCAGCAGCTCCAGCCCCCAGATTACGCTGCCTCACAC 514
QY 594 GGTCCAGTGCCTCAACCTCTCCATCTCCCATGCCACCTGCCATGGTGTGTATCCCGG 653
DB 515 CTTGCGATGCGCCACATCACCATTGAGCACCAGAAAGTGTGAGAACGCTACCCCGG 574
QY 654 GAGATCAGAGCAACATGGTGTGTGAGGGG---GCGTCCGGGGGAGGATGCTGCGCA 710
DB 575 CAACATCAGACACCATGGTGTGTGCCAGCGTGCAGGAAGGGGGCAAGGACTCTCGCCA 634
QY 711 GGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGAGTCTTCAAGGTCTGGTGTCTGGGG 770
DB 635 GGGTGACTCGGGGGCCCTCTGGTCTGTAAACAGTCTCTTCAAGGATATATCTCTGGGG 694
QY 771 GTCTGTGGGGCCCTGTGGACAGATGGCATCCCTGGAGTCTACACCTATATTGCAAGTA 830
DB 695 ---CCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAGAAAGTCTGCAATA 751
QY 831 TGTGACTGGATCCGGATGATCATGAGGAAACACTGACCTGTTTCTCAGCTTCACCCC 890
DB 752 TGTGACTGGATCCAGGAGACGATGAGAAACAATTAGACTGGACNTCACTCCGAANCC 811
QY 891 C 891
DB 812 C 812

RESULT 9
US-09-386-642-10
; Sequence 10, Application US/093866642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-10

Query Match 18.3%; Score 199.4; DB 4; Length 1052;
Best Local Similarity 59.6%; Pred. No. 2.3e-45;
Matches 395; Conservative 0; Mismatches 256; Indels 12; Gaps 3;
QY 212 ACTCAGCGCTGGCAGGTGGGGCTGTTTGGGGGACACGCTGCGCTGGGGGGTGTCC 271
DB 197 ACTCCAGCCCTGGCAGGAGCCCTGTTGGAAGACGCGGCTACTCTGTGGGGGAGCGC 256
QY 272 TTATTGACCACAGGTGGGTCTCTCACAGCGGTCTCACTGACGCGGAGCAGGTACTGGGTGC 331
DB 257 TCATGCCCGCCAGATGGTCTCTGACAGAGCCCACTGCTCAAGCCCCGCTACATAGTTC 316
QY 332 GCGTGGGGAAACAGCCTCAGCCAGTCACTGACGAGCAGATCCGGCAGCAGCGGT 391
DB 317 ACCTGGGGCAGCACAACTCTCCAGAAAGGAGGAGGCTGTGAGCAGACCCCGCAGCCACTG 376
QY 392 TCTCTGTGACCCATCCCGCTACCTGGGAGCTTCAGCAGAGCCACGAGCAC-----GACC 445
DB 377 AGTCTTCTCCCGACCCCGGCTTCAACACAGCCTCTCCCAACAAAGACCAACCGCAATGACA 436

Db 705 GGTGGCACTGGTCTTACAGTCCCTGTAATAAACCAGAAATACACAGTACGCGTG 764
QY 337 GGGGAACACAGCTCAGCAGCTGACTGAGCAGGAGAGATCGGCACAGCGGCTTCTCT 396
Db 765 GGAGACACACAGCTTACAGAAATAGATGGCCAGAGCAAGAAATACCTGTGGTTAGTCC 824
QY 397 GTGACCATCCCGGT--ACCTGGGAGCTCGACGAGCCACGAGCAACACTCCGCGTG 453
Db 825 ATCCACACCCCTGCTACACAGCAGCATGGAGGACCAACCATGATCTGATGCTT 884
QY 454 CTGCGCTGCGCTGCCCTCCGCTGCGCTGACGAGCGTTCAACCCCTGCCCTGCCCAAT 513
Db 885 CTTCAACTGCGTGACGAGCATCCCTGGGCTCCAAAGTGAAGCCCATCAGCCTGGCAGAT 944
QY 514 GACTGTGCAACCGCTGGCAGGAGTCCAGCTCTCAGGCTGGGGCATCACCACCA 573
Db 945 CATTCACCCAGCTGCGCCAGAGAGTGCACCGCTCTCAGGCTGGGGCACTGTACCACTCCC 1004
QY 574 CGGAACCCATCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCATGCCACC 633
Db 1005 CGAGAAATTTCTGACACTCTCACTGTGAGAAGTAAATCTTTCCCGAGAGAG 1064
QY 634 TGCCATGGTGTATCCCGGAGAAATCACGAGCAACATGTTGTGTCAGCGCGGCTCCG 693
Db 1065 TGTGAGGATGTTACCCGGGGCAGATCACAGATGGCATGCTGTGTCAGGCGAGCAAA 1124
QY 694 GGGCAGGATGCTCCAGGAGTGTCTGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAA 753
Db 1125 GGGGCTGACACGTCGCGAGGCGATCTGAGAGCCCCCTGTTGTGTTGATGGTGCACCTCCAG 1184
QY 754 GGTCTGTTGCTCCGGGCTCTGTTGGGGCCCTGTGGCAAGATGGCATCCCTGGAGTCTAC 813
Db 1185 GGCATCACATCTCTGGGCTC--AGACCCCTGTGGAGTCCGACAAACTTGGCGTCTAT 1241
QY 814 ACCTATTGCAAGTATGTGAGTCCGATCCGATGATCATGAGGAACAA 863
Db 1242 ACCAACATCTGCGCTACCTGGACTGGATCAAGAGATCATAGGCAGCAA 1291

RESULT 12

US-09-008-271A-19
; Sequence 19, Application US/09008271A
; Patent No. 6203979
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 16-Jan-1998
; PRIOR APPLICATION NUMBER: US/09/008,271A
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 994 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT27
; CLONE: 1798496
; SEQUENCE DESCRIPTION: SEQ ID NO: 19 :
US-09-008-271A-19

Query Match 17.4%; Score 190; DB 3; Length 994;
Best Local Similarity 56.2%; Pred. No. 8 7e-43;
Matches 399; Conservative 0; Mismatches 305; Indels 6; Gaps 2;
QY 157 GGGCTCAGCCAGGCGAGCCACACCGAAGATTTTCAATGGCAGCTGAGTGTGGCGTAACCTCA 216
Db 224 GGACACTCCAGGGCACAGGAGCAAGGTGCTGGGGGGTCATGAGTGCACACCCCATTCG 283
QY 217 CAGCGTGGCAGGTGGGCTGTTGAGGGCACACGCTCGGTCGGGGTGTCTTATT 276
Db 284 CAGCCTTGGCAGCGGCTTGTCCAGGGCCAGCAACTACTCTGTGGGGTGTCTTGTGA 343
QY 277 GACCAAGTGGGTCTCTCACAGCGGCTCACTGCAGCGGCGAGGATCTGCGGCTGCGCCTG 336
Db 344 GTGGCACTGGGTCTTACAGCTGCCACTGTAAACCCGAAATACACAGTACGCGCTG 403
QY 337 GGGGAACACAGCCTCAGCCAGCTCGACTGGAACGAGCAGATCCGGCAGCGGCTTCTCT 396
Db 404 GGAGACCAAGCTTACAGAAATAAGATGGCCAGAGCAAGAAATACCTGTGGTTCAGTCC 463
QY 397 GTGACCCATCCCGGT--ACCTGGGAGCTCGACGAGCGATGAGGAGCCACCAACATGATCTGATGCTT 523
Db 464 ATCCACACCCCTGCTTACAAAGCAGCGATGAGGAGCCACCAACATGATCTGATGCTT 523
QY 454 CTGCGCTCGCTGCGCTCGCGCTAACAGCAGCGTTCAACCCCTGCGCTGCGCTGCGCAAT 513
Db 524 CTTCAACTGGTGACCGAGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGAT 583
QY 514 GACTGTGCAACCGCTGGCAGCGAGTCCAGCTCTGAGGCTGGGGCATCACCACCAACCA 573
Db 584 CATTCACCCAGCTGGCCAGAAAGTGCACCGCTCTCAGGCTGGGGCACTGTCAACAGTCC 643
QY 574 CGGAACCCATCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCTCCCATGCCACC 633
Db 644 CGAGAAATTTCTGACACTCTCAACTGTGAGAAAGTAAATCTTTCCCGAGAGAG 703
QY 634 TGCCATGGTGTATCCCGGAGAAATCACAGCAACATGTTGTGTCAGCGCGGCTCCCG 693
Db 704 TGTGAGATGCTTACCCGGGGCAGATCACAGATGGCATGCTGTGTCAGGCGAGCAGCAAA 763
QY 694 GGGCAGGATGCTGCCAGGATGTTCTGGGGCCCCCTGTTGTGTTGGGGAGTCTTCAA 753
Db 764 GGGGCTGACACGTCGCGAGGGCGATTTCTGGAGGCCCTGTTGTGTTGATGCTGACCTCCAG 823
QY 754 GGTCTGCTGCTGCGGGTCTGTGGGGCCCTGTGGACAGAGTGGATCCTCGAGTCTAC 813
Db 824 GGCATCACATCTGGGGCTC--AGACCCCTGTGGAGGTCCGACAAACTGGCGTCTAT 880
QY 814 ACCTATTGCAAGTATGTGAGTCCGATCCGATCATGAGGAACAA 863
Db 881 ACCAACATCTGCGCTACCTGGACTGGATCAAGAGATCATAGGCAGCAA 930

RESULT 13

US-09-386-642-9

; Sequence 9, Application US/09386642

; Patent No. 6420157

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew

; APPLICANT: Qi, Jensen

; APPLICANT: Andrade-Gordon, Patricia

; TITLE OF INVENTION: Zymogen Activation System

; FILE REFERENCE: ORT-1028

; CURRENT APPLICATION NUMBER: US/09/386,642

; CURRENT FILING DATE: 1999-08-31

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 1049

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene

; OTHER INFORMATION: with homo sapien serine protease catalytic domain

US-09-386-642-9

Query Match 17.3%; Score 188.4; DB 4; Length 1049;

Best Local Similarity 57.4%; Pred. No. 2.5e-42;

Matches 380; Conservative 0; Mismatches 276; Indels 6; Gaps 2;

QY 212 ACTCAGCGGTGGCAGGTGGGGCTGTTTGGAGGACACAGCTGGCGTGGGGGTGTC 271

DB 197 ATTGGAGCTTGGCAGGGCGCTGTTCAGGGCCAGCACTCTCTGTGGCGGTGCC 256

QY 272 TTATTGACCACAGGTGGGTCTTACAGCGGCTCACTGACGGCGCAGCGTACTGGGTGC 331

DB 257 TTGTAGTGGCACTGGGTCTTACAGCTGCCACTGTAAAAACCGAATAACAGTAC 316

QY 332 GCGTGGGGAACACAGCTCAGCGCTCAGCTGACGACGAGTCCGGCAGCGGCT 391

DB 317 GCGTGGGACACAGCTTACAGTAATAAGATGGCCAGCAGAAATACCTGTGGTTC 376

QY 392 TCTGTGTACCATCCGGCT--ACCTGGAGCTTCGACGAGCCACGACGACGCTCC 448

DB 377 AGTCCATCCACACCTCTACACAGCAGCGATGTGGAGGCCACACACCATGATCTGA 436

QY 449 GGTGTGGGTGGCGCTGCCGTCCGGTAAACAGCAGCGTTCACCCCTGCCCTGTC 508

DB 437 TGCTTCTCAACTGCGTGACCGGCTCCCTGGGTCCAAAGTGAAGCCCATCAGCCTGG 496

QY 509 CCAATGACTGTCAACCGCTGGCAGCGTCCAGCTCAGCTGGGCGCATCACCAACC 568

DB 497 CAGTCAATGACCCAGCTGGCCAGAGTGACCGTCTCAGGCTGGGCGACTGTCAACA 556

QY 569 ACCACGGAACCAATCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTCCCATG 628

DB 557 GTCCCGGAGAGATTTCTTGACACTCTCAACTGTGACAGATTAATAATCTTTCCCA 616

QY 629 CACCTGCCATGTGTGTATCCCGGAGATCACAGACACATGTGTGTGAGGGGGG 688

DB 617 AGAAGTGTGAGGATGTCTACCGGGGCGAGATCAAGATGGCTGTGAGCGCAGCA 676

QY 689 TCCCGGGCAGGATGCTGTCAGGGTGAATCTGGGGGCCCTGTGTGTGGGGAGTCC 748

DB 677 GCMAAGGGGTGACAGTGGCCAGGGCGATCTGAGAGCCCTGTGTGTGTGTGTGAC 736

QY 749 TTCAGGCTGTGTGCTCTGGGGTCTGTGGGGCCCTGTGTGACAAAGATGCAATCCCTGAG 808

DB 737 TCCAGGGCATCATCTCTGGGGCTC--AGACCCCTGTGGGAGGTCCGACAAACCTGGCG 793

QY 809 TCTACACCTATTTTGAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAATGAC 868

DB 794 TCTATACCAACATCTGCGGTACCTGGACTGGATCAAGAGATCATAGGCAGCAAGGGCT 853

QY 869 CT 870

DB 854 CT 855

RESULT 14

US-09-070-526-1

; Sequence 1, Application US/09070526

; Patent No. 6100059

; GENERAL INFORMATION:

; APPLICANT: SOUTHAN, CHRISTOPHER

; APPLICANT: CLINKENBEARD, HELEN

; APPLICANT: BURGESS, NICOLA

; TITLE OF INVENTION: NO. 6100059el Compounds

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/070,526

; FILING DATE: 30-APR-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9711952.3

; FILING DATE: 9-JUN-1997

; APPLICATION NUMBER: EP 9730964.4

; FILING DATE: 1-DEC-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-30353

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 944 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-09-070-526-1

Query Match 16.9%; Score 184.8; DB 3; Length 944;

Best Local Similarity 55.6%; Pred. No. 2.3e-41;

Matches 395; Conservative 1; Mismatches 308; Indels 6; Gaps 2;

QY 157 GGGCTCAGCCAGGCGCCACACCGAAGATTTCAATGCCACTGAGTGTGGGGTAACCA 216

DB 174 GGACACTCCAGGCGCACAGAGGACAAAGGTGCTGGGGGTCTATGAGTGCAC 233

QY 217 CAGCGCTGGCAGGTGGGCTGTTTTCAGGGCCACCGCTCGCGTGGGGGGTGTCTTATT 276

DB 234 CAGCTTGGCAGGCGGCTTGTTCAGGGCCAGCAATCTCTGTGGCGGTGCTCTGTA 293

QY 277 GACCAAGGTGGTCTCTCAGCGGCTCATGTCAGCGGCGAGGAGTCTGGGTGGCGCTG 336

DB 294 GGTGGCAACTGGGTCTTACAGTCCCACTGTAAAAAACCGAAATACACAGTACGCTG 353

QY 337 GGGGACACAGCTCAGCCAGCTCCAGTGGACCGGAGAGATCCGGCACAGCGGCTCTCT 396

DB 354 GGAGACACAGCTTACAGATTAAGATGGCCAGAGCAAGAAATACCTGTGTTTCACTCC 413

QY 397 GTGACCCATCCGGCT---ACCTGGAGGCTCGAGCGCCAGCGAGCACGACCTCGGGCTG 453

DB 414 ATCCACACCCCTGCTATATACAGCGGATGTGGAGGACCAACCATGATCTGATGCTT 473

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 05:53:50 ; Search time 4625 Seconds
(without alignments)
10224.270 Million cell updates/sec

Title: US-10-015-385A-193

Perfect score: 1091

Sequence: 1 caagcaggcatccctctgg.....aggggcaaaaaaaaaa 1091

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: gb.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1091	100.0	1091	6	AX697125	Sequence
2	1091	100.0	1091	9	AY358524	AY358524 Homo sapi
3	498	45.6	811	9	HSMB08731	BS648580 Homo sapi
4	487	44.6	619	6	AR263961	Sequence
5	440	40.3	840	6	AR337568	Sequence
6	440	40.3	840	6	BD130047	Human ser
7	370	33.9	9120	9	AF164623	Homo sapi
8	370	33.9	11820	9	AF135025	Homo sapi
9	370	33.9	132323	9	AC011473	Homo sapi
10	370	33.9	230000	9	AF243527	Homo sapi
11	249	22.8	142334	2	AC073185	Homo sapi
12	227	20.8	405	6	AX070938	Sequence
13	217	19.9	174724	2	AC140096	Pan trogl
14	217	19.9	200792	2	AC130782	Pan trogl
15	127	11.6	142334	2	AC073185	Homo sapi
16	75	6.9	176647	2	AC130188	Papio anu
17	26	2.4	938	4	AB100595	Canis fam
18	25	2.3	25	6	AX697374	Sequence
19	24	2.2	24	6	AX697375	Sequence
20	24	2.2	110000	2	AC119867_2	Continuation (3 of
21	24	2.2	154879	9	AL450104	Human DNA
22	24	2.2	159391	2	AC027113	Homo sapi
23	24	2.2	167302	9	AC099756	Homo sapi
24	24	2.2	175572	2	AC025349	Homo sapi
25	24	2.2	184895	5	AL844197	Zebrafish
26	24	2.2	186971	9	AP003501	Homo sapi
27	24	2.2	226890	2	AC135541	Rattus no
28	24	2.2	228413	2	AC127853	Rattus no
29	24	2.2	243655	2	AC099172	Rattus no
30	23	2.1	578	11	G78591	S209P6196PD
31	23	2.1	786	6	AX375740	Sequence
32	23	2.1	944	6	AX375744	Sequence
33	23	2.1	954	4	PIGELAI	Sus scrofa
34	23	2.1	959	4	SSELS1	Sus scrofa
35	23	2.1	970	6	E00424	Pig elastas
36	23	2.1	1020	6	AX823252	Sequence
37	23	2.1	1130	6	AR256996	Sequence
38	23	2.1	1606	6	AX676264	Sequence
39	23	2.1	1613	6	AR256990	Sequence
40	23	2.1	1718	9	AF536382	Homo sapi
41	23	2.1	1718	9	BC062334	Homo sapi
42	23	2.1	1755	9	BC036846	Homo sapi
43	23	2.1	110000	2	AC079517_1	Continuation (2 of
44	23	2.1	159242	8	AP002902	Oryza sat
45	23	2.1	159440	8	AP002900	Oryza sat

ALIGNMENTS

RESULT 1
AX697125
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX697125
Sequence 193 from Patent WO0078961.
AX697125
AX697125.1 GI:29498087
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
1
Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Gadowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,

AX697125
Sequence 193 from Patent WO0078961.
AX697125
AX697125.1 GI:29498087
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
1
Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Gadowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,

AX697125
Sequence 193 from Patent WO0078961.
AX697125
AX697125.1 GI:29498087
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
1
Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Gadowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,

Query Match 100.0%; Score 1091; DB 9; Length 1091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCAGGTCTATCCCTTGGTGACCTTCAAGAGAGACAGAGAGGCGAGAGTGGGGGGC 60
DB 1 CAAGCAGGTCTATCCCTTGGTGACCTTCAAGAGAGACAGAGAGGCGAGAGTGGGGGGC 60

QY 61 ACAGGGAAGAGGTGACCTCTGAGATTCCCTTTTCCCCAGACACTTTGGAGAGTACCCACC 120
DB 61 ACAGGGAAGAGGTGACCTCTGAGATTCCCTTTTCCCCAGACACTTTGGAGAGTACCCACC 120

QY 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTTGTTCTTTGGGCTCAGCCAGGAGCCACACCG 180
DB 121 ATGGGGCTCAGCATCTTTTGGTCTCTGTTGTTCTTTGGGCTCAGCCAGGAGCCACACCG 180

QY 181 AAGATTTTCAATGGCACTGAGTGGGGCTTAATCTACAGCCGTGGAGTGGGGCTGTTT 240
DB 181 AAGATTTTCAATGGCACTGAGTGGGGCTTAATCTACAGCCGTGGAGTGGGGCTGTTT 240

QY 241 GAGGGCACCAGCGCTGGCTGCGGGGTGCTTTATTGACACAGAGTGGGTCTCTCACAGCG 300
DB 241 GAGGGCACCAGCGCTGGCTGCGGGGTGCTTTATTGACACAGAGTGGGTCTCTCACAGCG 300

QY 301 GCTCACTGAGCGGACAGAGTACTGGGTGCGCTGGGGGGAACACAGCTCAGCCAGCTC 360
DB 301 GCTCACTGAGCGGACAGAGTACTGGGTGCGCTGGGGGGAACACAGCTCAGCCAGCTC 360

QY 361 GACTGACCGCAGCATCGGCACAGCGGCTTCTGTGACCCATCCCGGCTACCTGGGA 420
DB 361 GACTGACCGCAGCATCGGCACAGCGGCTTCTGTGACCCATCCCGGCTACCTGGGA 420

QY 421 GCTCGACAGCAGCAGCAGCAGCTCCGGCTGCTGCGGCTGCGCTGCGGCTGCGCGTA 480
DB 421 GCTCGACAGCAGCAGCAGCAGCTCCGGCTGCTGCGGCTGCGCTGCGGCTGCGCGTA 480

QY 481 ACCAGCAGGTTCAACCCCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 540
DB 481 ACCAGCAGGTTCAACCCCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 540

QY 541 CAGCTCTCAGGCTGGGGCATACCAACACACCCAGGAGCCATTCCTCGGATCTGCTCCAG 600
DB 541 CAGCTCTCAGGCTGGGGCATACCAACACACCCAGGAGCCATTCCTCGGATCTGCTCCAG 600

QY 601 TGCCTCAACCTCTCCATGCTTCCCATGCGCATGCGGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 TGCCTCAACCTCTCCATGCTTCCCATGCGCATGCGGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 661 ACGAGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 ACGAGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 721 GGGGGCCCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 GGGGGCCCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 781 CCTGTGACAGAGTGGCATCTCTGAGTCTCACTATATTTGCAAGTATGTAAGTGG 840
DB 781 CCTGTGACAGAGTGGCATCTCTGAGTCTCACTATATTTGCAAGTATGTAAGTGG 840

QY 841 ATCCGGATGATCATGAGGAACACTGACCTGTTCTTCCACCTCCACCCCGCCCTTAA 900
DB 841 ATCCGGATGATCATGAGGAACACTGACCTGTTCTTCCACCTCCACCCCGCCCTTAA 900

QY 901 CTTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCTCTCCATCACTTCCCTAGCTCC 960
DB 901 CTTGGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCTCTCCATCACTTCCCTAGCTCC 960

QY 961 ACTCTGTGGGCTGGGAACCTTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCA 1020
DB 961 ACTCTGTGGGCTGGGAACCTTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCA 1020

QY 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATGAAGAGGGGGCANA 1080

DB 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATGAAGAGGGGGCANA 1080

QY 1081 AAAAAAAAAAAAA 1091
DB 1081 AAAAAAAAAAAAA 1091

RESULT 3
HSM808731
LOCUS HSM808731 811 bp mRNA linear PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686H1078 (from clone DKFZp686H1078).
ACCESSION BX648580
VERSION BX648580.1 GI:34367742
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Bloeker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobe, G., Han, M. and Wiemann, S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686H1078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
Location/Qualifiers
1..811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686H1078"
/tissue_type="human endometrium carcinoma cell line"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polyA_signal 702..707
polyA_site 723

ORIGIN
Query Match 45.6%; Score 498; DB 9; Length 811;
Best Local Similarity 100.0%; Pred. No. 1.4e-284; Mismatches 0; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 578 ACCCATCCCGGATCTGCTCCAGTGCCTCAACCTTCATCGTCTCCATGCCACCTGCC 637
DB 226 ACCCATCCCGGATCTGCTCCAGTGCCTCAACCTTCATCGTCTCCATGCCACCTGCC 285

QY 638 ATGGTGTATATCCCGGAGAAATCAGAGCAACATGGTGTGTGCGAGCGCGTCCCGGGC 697
DB 286 ATGGTGTATATCCCGGAGAAATCAGAGCAACATGGTGTGTGCGAGCGCGTCCCGGGC 345

QY 698 AGGATGCTGCGCAGGGTGAATTCGGGGGCCCTTGGTGTGTGGGGAGTCTCTCAAGTCT 757
DB 346 AGGATGCTGCCAGGGTGAATTCGGGGGCCCTTGGTGTGTGGGGAGTCTCTCAAGTCT 405

QY 758 TGGTGTCTCTGGGGTCTCTGGGGGCCCTTGTGAAGATGGCATCCCTGGAGTCTACACT 817
DB 406 TGGTGTCTCTGGGGTCTCTGGGGGCCCTTGTGAAGATGGCATCCCTGGAGTCTACACT 465

QY 818 ATATTGCAAGTATGACCTGGATCCGGATGATCATGAGCAACACTGACCTGTTTCT 877
DB 466 ATATTGCAAGTATGACCTGGATCCGGATGATCATGAGCAACACTGACCTGTTTCT 525


```
KEYWORDS      JP 2002502600-A/3.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 840).
AUTHORS      Ruben,S.M. and N.J.J.
TITLE        Human serine protease and serpin polypeptide
JOURNAL      HUMAN GENOME SCIENCES INC
              Patent: JP 2002502600-A 3 29-JAN-2002;
              HUMAN GENOME SCIENCES INC
COMMENT      OS Homo sapiens (human)
              PN JP 2002502600-A/3
              PD 29-JAN-2002
              PF 04-FEB-1999 JP 2000530597
              PR 06-FEB-1998 US 60/073961
              PI STEVEN M RUBEN, JIAN NI
              PC
              C12N9/64, A61K38/48, A61P7/02, A61P7/04, A61P7/06, A61P29/00, A61P31/
              00,
              PC A61P35/00, A61P37/06, C07K14/81, C07K16/40, C12N1/15, C12N1/19, PC
              C12N1/21,
              PC C12N5/10, C12N15/09, C12O1/37, A61K37/547, C12N5/00, C12N15/00 CC
              Human serine protease and serpin polypeptide
              PH Key Location/Qualifiers
              FT CDS (115)..(603).
              Location/Qualifiers
              1..840
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"

FEATURES             source
    source
    Join(2313..2398,4189..4263,5061..5217,5545..5810,
    5627..6763,7158..7622)
    gene
    mRNA
    CDS
    ORIGIN
    Query Match      40.3%; Score 440; DB 6; Length 840;
    Best Local Similarity 99.8%; Pred. No. 5.7e-250;
    Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
    QY 595 CCGGATCTGCTCCAGTGCCTCAACTCTCCATCGTCTCCCATGCGACCTGCATGGTGT 644
    Db      |||
    QY 321 CCGGATCTGCTCCAGTGCCTCAACTCTCCATCGTCTCCCATGCGACCTGCATGGTGT 380
    Db      |||
    QY 645 GTATCCCGGGAGAAATCACCAGCAACATGTGTGTGTCAGCGCGCGCTCCCGGGCAGGATGC 704
    Db      |||
    QY 381 GTATCCCGGGAGAAATCACCAGCAACATGTGTGTGTCAGCGCGCGCTCCCGGGCAGGATGC 440
    Db      |||
    QY 705 CTGCCAGGATGATCTGCGGGGCCCTGTGTGTGTGCGGGAGTCTTCAAGGTCTGTGTGTC 764
    Db      |||
    QY 441 CTGCCAGGATGATCTGCGGGGCCCTGTGTGTGTGCGGGAGTCTTCAAGGTCTGTGTGTC 500
    Db      |||
    QY 765 CTGGGGGTCTGTGGGGCCCTGTGGACAGATGCGCATCCCTGGAGTCTACACCTATATTG 824
    Db      |||
    QY 501 CTGGGGGTCTGTGGGGCCCTGTGGACAGATGCGCATCCCTGGAGTCTACACCTATATTG 560
    Db      |||
    QY 825 CAAATATGTGGAAGTGGATCCGGATGATCATAGGAACAACACTGTTTCTCCACCTC 884
    Db      |||
    QY 561 CAAATATGTGGAAGTGGATCCGGATGATCATAGGAACAACACTGTTTCTCCACCTC 620
    Db      |||
    QY 885 CACCCGCCACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGACACCAATCTCTCCAT 944
    Db      |||
    QY 621 CACCCGCCACCCCTTAACCTTGGGTACCCCTCTGGCCCTCAGACACCAATCTCTCCAT 680
    Db      |||
    QY 945 CACTTCCCTTACTCCACTCTGTGTGGCTGGGAACTCTTGGAACTTTAACTCTGCA 1004
    Db      |||
    QY 681 CACTTCCCTTACTCCACTCTGTGTGGCTGGGAACTCTTGGAACTTTAACTCTGCA 740
    Db      |||
    QY 1005 GCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATAA 1064
    Db      |||
    QY 741 GCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATAA 800
    Db      |||
    QY 1065 ATGAAGGAGGG 1075
    Db      |||
    QY 801 ATGAAGGAGGG 811
    Db      |||
```



```
repeat_region      /rpt_family=" (TTTG)n"
repeat_region      complement(11133..11411)
repeat_region      /rpt_family="AluSg"
repeat_region      11412..11515
repeat_region      /rpt_family="MIR"
repeat_region      11680..11979
repeat_region      /rpt_family="AluSx"
repeat_region      complement(11989..12069)
repeat_region      /rpt_family="MIR"
repeat_region      complement(12070..12390)
repeat_region      /rpt_family="AluJo"
repeat_region      complement(12391..12498)
repeat_region      /rpt_family="MIR"
repeat_region      12503..12606
repeat_region      /rpt_family="L2"
repeat_region      complement(12694..12976)
repeat_region      /rpt_family="L2"
repeat_region      12993..13101
repeat_region      /rpt_family="L2"
repeat_region      13287..13393
repeat_region      /rpt_family="L2"
repeat_region      complement(13507..13804)
repeat_region      /rpt_family="AluS4"
repeat_region      complement(14351..14379)
repeat_region      /gene="K1K10"
repeat_region      /note="Kallikrein 10; synonyms: NES1, PRSSL1"
repeat_region      complement(join(14351..15033,15498..15631,15963..16237,
repeat_region      17191..17371,19124..19220,19708..19779))
repeat_region      /gene="K1K10"
repeat_region      /product="K1K10 mRNA"
repeat_region      14384..14530
repeat_region      /standard_name="stSG30247"
repeat_region      /note="RH48452, WTAF-3662-STs"
repeat_region      complement(14432..14604)
repeat_region      /rpt_family="MIR"
repeat_region      complement(join(14881..15033,15498..15631,15963..16237,
repeat_region      17191..17371,19124..19220))
repeat_region      /gene="K1K10"
repeat_region      /note="protease, serine-like, 1"
repeat_region      /codon_start=1
repeat_region      /product="kallikrein 10"
repeat_region      /protein_id="AAG23256.1"
repeat_region      /db_xref="GI:10799395"
repeat_region      /translation="MRPHLHLSAAGARALAKLPLIMAAEALPQNDTRL
repeat_region      PEAYGSPCARQSPQVSLFNGLSFHCAGVLVDQSVLTAHCKNPLWARVGDHLL
repeat_region      LLOEQLRTRVVPKHYOGSGPILPRRDEHLLMLKLARPVVLPGRVALQLEP
repeat_region      RCAOPGQCOVAGMGTAAARVKNKGLTCSITILSPKECEVPYGVVNMICAGL
repeat_region      DRGQDPCQSDSGGLVCDTELQGLSNGVYFCGSAQHPAVYTIQIKTMSWINKVIRSN
repeat_region      "
repeat_region      15245..15420
repeat_region      /rpt_family=" (TTGGGG)n"
repeat_region      15715..15967
repeat_region      /rpt_family="MSR1"
repeat_region      complement(16473..16818)
repeat_region      /rpt_family="AluS4"
repeat_region      17787..17912
repeat_region      /rpt_family="AluJo"
repeat_region      17913..17932
repeat_region      /rpt_family=" (CAA)n"

Query Match      33.9%; Score 370; DB 9; Length 132323;
Best Local Similarity 100.0%; Pred. No. 3.5e-208;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      708 CCAGGGTGAATCTGGGGGCCCTGGTGTGTGGGGAGTCTCTTCAAGGTCTGGTCTCTG 767
Db      29542 CCAGGGTGAATCTGGGGGCCCTGGTGTGTGGGGAGTCTCTTCAAGGTCTGGTCTCTG 29483

QY      768 GGGGTCTGGGGCCCTGTGACAAAGATGGCATCCCTGGAGTCTACACCTATATTGCAA 827
Db      29482 GGGGTCTGGGGCCCTGTGACAAAGATGGCATCCCTGGAGTCTACACCTATATTGCAA 29423

QY      828 GTATGTGGACTGGATCCGGATGATCATGAGGAACACTGACCTGTTCTTCTCCACCTCCAC 887
```

```
Db      29422 GTATGTGGACTGGATCCGGATGATCATGAGGAACACTGACCTGTTCTTCTCCACCTCCAC 29363
QY      888 CCCACACCCCTTAATCTGGGTACCCCTCTGGCCCTCAGAGCACCACCAATATCTCTCCATCAC 947
Db      29362 CCCACACCCCTTAATCTGGGTACCCCTCTGGCCCTCAGAGCACCACCAATATCTCTCCATCAC 29303
QY      948 TTCCCTAGCTCCACTCTCTGTGTGGCCCTGGGAACCTTCTTGGAACTTTAACTCTCCGAGCC 1007
Db      29302 TTCCCTAGCTCCACTCTCTGTGTGGCCCTGGGAACCTTCTTGGAACTTTAACTCTCCGAGCC 29243
QY      1008 CTTCTAAGACCCAGAGCGGGGTGAGAGAACTGTGCATAGTCTGGATATAATTAATG 1067
Db      29242 CTTCTAAGACCCAGAGCGGGGTGAGAGAACTGTGCATAGTCTGGATATAATTAATG 29183
QY      1068 AAGGAGGGGC 1077
Db      29182 AAGGAGGGGC 29173

RESULT 10
AF243527/c
LOCUS      AF243527      230000 bp      DNA      linear      PRI 21-NOV-2000
DEFINITION Homo sapiens serine protease gene cluster, complete sequence.
ACCESSION      AF243527
VERSION      AF243527.1 GI:11244757
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 230000)
AUTHORS      Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepfer, B. and Wang, K.
TITLE      Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL      Gene 257 (1), 119-130 (2000)
MEDLINE      20510030
PubMed      11054574
REFERENCE      2 (bases 1 to 230000)
AUTHORS      Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepfer, B. and Wang, K.
TITLE      Direct Submission
JOURNAL      Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
source
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/complement(<6963..>11517)
/gene="KLK1"
9482..9641,11472..>11517)
/gene="KLK1"
/product="renal kallikrein"
complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
/note="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG3353.1"
/db_xref="GI:11244758"
/translati="MMFLVLCALSLGGTGAAPPIQSIRVGVWECEHSPQWAAALYH
FSTFCGGILVHRQWLTAHCISDNVQLWGRHNLFDDENTAQFVHVVSFPHGFN
MSLLEHTRQADEYSDHMLLRLTEPADTITDAVKVVELPTEPVGSTCLASGFS
IERENSPFDDLCQVDLKLILPNDEKKAHVQKVTDFMLCVGLEKNDTCVGDSGGFL
MCDVPGQVTSWGIYPCGIPKPSVAVRVLVSUYVKWIEDTIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
mRNA
```

```

CDS
/complement (join (13552, .13704, 14377, .14530, 14678, .14917,
15416, .15569, 19204, .19246))
/notes="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MWLLLTLSFLASTAAQDQKLEGEBCAPHSPQWVALYERGR
FNCAGLISPHWVLSAAHCSRMVRLGHNLRKDPGRLRTSVIHPHYEARS
HRNDIMLLRLVQARLNPQVRPAVLPRCPHGEACVSGWGLVPLSPVSLDILHC
ANTSISDTSCKSKYRGLNTWVCAGARGAESCEGSGGLVCGGILQGVSMGD
VPCDNTTKPGVYTKVHYLEWIRETKRN"
<42595, .>47769
/gene="KLK3"
mRNA
join (<42595, .42640, 43880, .44039, 45669, .45955, 46099, .46235,
47614, .>47769)
/gene="KLK3"
CDS
/product="prostate specific antigen"
/join (42595, .42640, 43880, .44039, 45669, .45955, 46099, .46235,
47614, .>47769)
/notes="KLK3"
/codon_start=1
/product="prostate specific antigen"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translation="MWVPVFLTSLVTWIGAAPLILSRIVGWCEKHSOPWQVIVAS
RGLACGVVLHPQWVLTAAHCRNKSIVLLGRHSILPHEDTGVQVSHSFHPLYD
MSLLKHQSLRPDESDSHDMLLRLSPAEITDAVQMDLPTQBPALGTTCYASGWSGI
EPEFETPKLQCVLDHIVSNDVCAQHPQKRTKMLCAGRTWGKSGTSGDGGGLV
CNGVLQITSWGSEPCALPERSLYTKVHYRWKIDTIANP"
<61139, .>66229
/gene="KLK2"
mRNA
join (<61139, .61184, 62391, .62550, 64142, .64428, 64542, .64678,
66074, .>66229)
/gene="KLK2"
CDS
/product="glandular kallikrein"
/join (61139, .61184, 62391, .62550, 64142, .64428, 64542, .64678,
66074, .>66229)
/notes="KLK2"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MWDLVLSIALSVGCTGAVPLIOSRIQVGVGWECEKHSQPMQVAVS
HGWAHCGVLVHPQWVLTAAHCLKNSQVWLGRHNLPEPDTQQRYPVSHSPHPLYN
MSLLKHQSLRPDESDSHDMLLRLSPAEITDAVQMDLPTQBPALGTTCYASGWSGI
EPEFETPKLQCVLSHLSDNMCARAYSEKVFMLCAGLWTGKDTCCGDSGGGLV
CNGVLQITSWGSEPCALPEKPAVYTKVHYRWKIDTIANP"
complement (join (<94602, .94754, 96027, .96163, 96247, .96497,
96919, .97081, 98345, .>98405))
/product="serine protease"
/complement (join (94602, 94754, 96027, .96163, 96247, .96497,
96919, .97081, 98345, .>98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="WATAGNPMGWFLGYLLVAGSLVSGSCQIINGDCSPHQPW
QALVWENELFCGVVLHPQWVLSAAHCFQNSYITIGLHLSLEADQPSQWVEASLS
VRHPYENRLLANDMLIKLDRSESDTSIRISASQCTPAGNSCLVSGWGLANGR
MTVLQCVNVSVEEVSKLYDPLYPHPSMFCAGGHDQKDCNGDSGGGLICNGYLQ
GLVFGKAPCGQGVGVVTLCKFTWETKTQAS"
complement (<131301, .>140303)
/gene="KLK5"
/notes="synonym: SCCE"
complement (join (<131301, .131456, 136310, .136443,
136529, .136785, 137525, .137690, 140255, .>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement (join (131301, .131456, 136310, .136443,

```

```

136529, .136785, 137525, .137690, 140255, .140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWVLCALITALLGVTDARSDSSSLINGSDDCMHTQWQAAL
LLRPNQLYCGAVLVHFQWMLTAACHCKRVFRVLGHYSLSPVYESGQPMFGVKSIPIH
PGVSHFGHNDMLIKLNRRIPTKVRPFINVSHCPASGAKTCLVSGWGTTKSPQVHE
FKYLQINTSLVLSQKCEADYPRQIDITMFCAGDKAGRDCSQDSDGPGVVCNGSLQGL
VSMGDYPCARPNRPVGVYTNLCKFTWETKTQAS"
complement (join (<146834, .146986, 149628, .149764,
151186, .151433, 155052, .155208, 155948, .>155987))
/product="protease M"
complement (join (146834, .146986, 149628, .149764,
151186, .151433, 155052, .155208, 155948, .155987))
/notes="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKKLMVLSLIAAAWAEQKLVHGGPCDKTSHPYQAALYTSGH
LTCGVLIHPLWLTAAHCKKNLQVFLGKNLRQRESSQSSVVRVILHPYDAAS
HDQDIMLLRLARPAKLSLIQPLERDCSANTSCHILGWGTADGDRPDTIQCAVI
HLVSRCECHAVPGQITONMLCAGDEKYGKDCSQDSDGGLVCGDHLRGLVSGNIFC
GSKEKPGVYTNVCRVTNWTQKTQAK"
complement (<165420, .>170283)
/gene="KLK7"
/notes="synonym: SCCE"
complement (join (<165420, .165575, 167672, .167808,
168124, .168371, 169651, .169798, 170211, .>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement (join (165420, .165575, 167672, .167808,
168124, .168371, 169651, .169798, 170211, .170283))
/notes="KLK7"
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MARSLILPIQLLSLALETAGEEAQGGKIIDGAPCARGSHPWQ
VALLSGNQLHCGVLNVRWLTAAHCKKNVTVHLGSDTLGDRRAORIKASKSRPH
GYSTOYVNDMLKNSQARLSVMVKVLRPSCEPPCTCTVSGWGTTSPTDFTFP
SLDMCVDVKLJSPQDCTKYDKLLENSMLCAGIPDSKKNACNGSDGGLVCRGTLQGL
VSWGTFPCQPDGVTYTVCAFTKWINDTMKHR"
complement (join (<183943, .184098, 185635, .185768,
187865, .188127, 188293, .188452, 188967, .>189036))
/product="neuropsin"
complement (join (183943, .184098, 185635, .185768,
187865, .188127, 188293, .188452, 188967, .189036))
/notes="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRPRPRAKTMWELLILGGAWAGHRAQERKVLGGHCQPHSQ
PMQAALFQGOQLLCGVLVGNWVLTAAHCKKPKTYVRLGDSHLQKDGPEQEIPIVQ
SIPHCYNSDVEDHNDMLLQLRDQASLSGSKVKPISLADHCTQPGQKCTVSGWGT
TSPFRFPDNLCAEVKIPQKKCEDAYPGQITDGMVCAAGSKSGADTCQDSDGGLVC
DGAQGITSWGSDPCGRDKPGVYTNICRYLDWIKKIISGK"
complement (join (<190980, .191129, 191573, .191709,
194324, .194589, 197048, .197204, 197370, .>197412))
/product="kallikrein-like 3"
complement (join (190980, .191129, 191573, .191709,
194324, .194589, 197048, .197204, 197370, .197412))
/notes="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translation="MKLLGALLSLLAGHWADTRAIABECRNSQPWQAGLPHLT
RLFCGATLISDRWLTAAHCKKPKVWLVRLEHLLWKWEGPQLFRVTVFFHPGPNKD

```

LSANDHDDIMLILPQARLSPAVPLNLSQTCVSPGMQCLISGWNVSPPKALFPV
TLQCANISILNKUICHWAYPHGISDMLCAGLWEGRGSCQDGGPLVNGTLAGVV
SGGAEPSPRRPRAVYTSVCHYLDWIOBEMEN"

Query Match 33.9%; Score 370; DB 9; Length 230000;
Best Local Similarity 100.0%; Pred. No. 3.6e-208; Mismatches 0; Indels 0; Gaps 0;
Matches 370; Conservative 0

QY	708	CCAGGGTGATCTGGGGGCCCTGGTGTGTGGGGAGTCTCTCAAGTCTGGTCTGCTG	767
Db	217325	CCAGGGTGATCTGGGGGCCCTGGTGTGTGGGGAGTCTCTCAAGTCTGGTCTGCTG	217266
QY	768	GGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCTGGAGTCTACACTATATTGCA	827
Db	217265	GGGGTCTGTGGGGCCCTGTGGACAGATGGCATCCTGGAGTCTACACTATATTGCA	217206
QY	828	GTATGTGACTGGATCCGGATGATCATGAGGAACAACCTGCTTCTCCACCTCCAC	887
Db	217205	GTATGTGACTGGATCCGGATGATCATGAGGAACAACCTGCTTCTCCACCTCCAC	217146
QY	888	CCCGACCCCTTAACCTGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCTCCATC	947
Db	217145	CCCGACCCCTTAACCTGGTACCCCTCTGGCCCTCAGAGCACCAATATCTCTCCATC	217086
QY	948	TTCCCTAGCTCCACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTGCGACCC	1007
Db	217085	TTCCCTAGCTCCACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTGCGACCC	217026
QY	1008	CTTCTAAGACCCACGAGCGGGTGAGAGTGTGCAATAGTCTGGAATAATAATAATG	1067
Db	217025	CTTCTAAGACCCACGAGCGGGTGAGAGTGTGCAATAGTCTGGAATAATAATAATG	216966
QY	1068	AAGGAGGGGC 1077	
Db	216965	AAGGAGGGGC 216956	

RESULT 11
AC073185/c
LOCUS AC073185 142334 bp DNA linear HTG 17-AUG-2000
DEFINITION Homo sapiens chromosome 19 clone RP11-10111, WORKING DRAFT
SEQUENCE, 32 unordered pieces.
ACCESSION AC073185
VERSION AC073185.4 GI:9838034
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
The sequence of Homo sapiens clone
1 (bases 1 to 142334)
2 (bases 1 to 142334)
Waterston,R.H.
Direct Submission
Submitted (10-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 17, 2000 this sequence version replaced gi:9653152.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH001011
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319

Consensus quality: bases at least Q40
Consensus quality: bases at least Q30
Consensus quality: bases at least Q20
Insert size: 135000; agarose-fp
Insert size: 139234; sum-of-contigs
Quality coverage: 3.86 in Q20 bases; agarose-fp
Quality coverage: 3.87 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	1452: contig of 1452 bp in length
*	1453	1552: gap of unknown length
*	1553	2737: contig of 1185 bp in length
*	2738	2837: gap of unknown length
*	2838	4665: contig of 1828 bp in length
*	4666	4765: gap of unknown length
*	4766	6054: contig of 1289 bp in length
*	6055	6154: gap of unknown length
*	6155	7791: contig of 1637 bp in length
*	7792	7891: gap of unknown length
*	7892	10028: contig of 2135 bp in length
*	10027	10126: gap of unknown length
*	10127	11895: contig of 1769 bp in length
*	11896	11895: gap of unknown length
*	11996	15070: contig of 3075 bp in length
*	15071	15170: gap of unknown length
*	15171	18478: contig of 3306 bp in length
*	18477	18576: gap of unknown length
*	18577	21537: contig of 2961 bp in length
*	21538	21637: gap of unknown length
*	21638	24625: contig of 2988 bp in length
*	24626	24725: gap of unknown length
*	24726	27583: contig of 2858 bp in length
*	27584	27684: gap of unknown length
*	27684	30892: contig of 3209 bp in length
*	30893	30992: gap of unknown length
*	30993	35209: contig of 4217 bp in length
*	35210	35309: gap of unknown length
*	35310	39261: contig of 3952 bp in length
*	39262	39361: gap of unknown length
*	39362	42356: contig of 2995 bp in length
*	42357	42456: gap of unknown length
*	42457	46421: contig of 3965 bp in length
*	46422	46521: gap of unknown length
*	46522	51157: contig of 4636 bp in length
*	51158	51257: gap of unknown length
*	51258	57416: contig of 6159 bp in length
*	57417	57516: gap of unknown length
*	57517	61710: contig of 4194 bp in length
*	61711	61810: gap of unknown length
*	61811	66249: contig of 4439 bp in length
*	66250	66349: gap of unknown length
*	66350	71980: contig of 5631 bp in length
*	71981	72080: gap of unknown length
*	72081	77454: contig of 5374 bp in length
*	77455	77554: gap of unknown length
*	77555	84318: contig of 6764 bp in length
*	84319	84418: gap of unknown length
*	84419	88649: contig of 4231 bp in length
*	88650	88749: gap of unknown length
*	88750	94419: contig of 5670 bp in length
*	94420	94519: gap of unknown length
*	94520	100165: contig of 5646 bp in length
*	100166	100265: gap of unknown length
*	100266	107334: contig of 7069 bp in length
*	107335	107434: gap of unknown length
*	107435	115093: contig of 7659 bp in length

```
* 115094 115193: gap of unknown length
* 115194 122023: contig of 6830 bp in length
* 122024 122123: gap of unknown length
* 122124 132649: contig of 10526 bp in length
* 132650 132749: gap of unknown length
* 132750 142334: contig of 9585 bp in length.
```

FEATURES

```
Location/Qualifiers
source
  1. .142334
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="19"
    /clone="RP11-10111"
  1. .1452
    /note="assembly_name:Contig26"
  1553. .2737
    /note="assembly_name:Contig27"
  2838. .4665
    /note="assembly_name:Contig28"
  4766. .6054
    /note="assembly_name:Contig29"
  6155. .7791
    /note="assembly_name:Contig30"
  7822. .10026
    /note="assembly_name:Contig31"
  10127. .11895
    /note="assembly_name:Contig33"
  11996. .15070
    /note="assembly_name:Contig34"
  15171. .18476
    /note="assembly_name:Contig35"
  18577. .21537
    /note="assembly_name:Contig36"
  21638. .24625
    /note="assembly_name:Contig37"
  24726. .27583
    /note="assembly_name:Contig38"
  27684. .30892
    /note="assembly_name:Contig39"
  30993. .35209
    /note="assembly_name:Contig40"
  35310. .39263
    /note="assembly_name:Contig41"
  39362. .42356
    /note="assembly_name:Contig42"
  42457. .46421
    /note="assembly_name:Contig43"
  46522. .51157
    /note="assembly_name:Contig44"
  51258. .57416
    /note="assembly_name:Contig45"
  57517. .61710
    /note="assembly_name:Contig46"
  61811. .66249
    /note="assembly_name:Contig47"
  66350. .71980
    /note="assembly_name:Contig48"
  72081. .77454
    /note="assembly_name:Contig49"
  77555. .84318
    /note="assembly_name:Contig50"
  84419. .88649
    /note="assembly_name:Contig51"
  88750. .94419
    /note="assembly_name:Contig52"
  94520. .100165
    /note="assembly_name:Contig53"
  100266. .107334
    /note="assembly_name:Contig54"
  107435. .115093
    /note="assembly_name:Contig55"
  115194. .122023
    /note="assembly_name:Contig56"
```

```
misc_feature 122124. .132649
              /note="assembly_name:Contig57
              clone_end:17
              vector_side:right"
misc_feature 132750. .142334
              /note="assembly_name:Contig58"

ORIGIN
Query Match      22.8%; Score 249; DB 2; Length 142334;
Best Local Similarity 99.7%; Pred. No. 5.3e-136;
Matches 369; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 708 CCAGGGTGATTTCTGGGGGCCCCCTGGTGTGTGGGGGAGTCCTTCAAGGTCTGGTGTCTG 767
Db 10984 CCAGGGTGATTTCTGGGGGCCCCCTGGTGTGTGGGGGAGTCCTTCAAGGTCTGGTGTCTG 10925
QY 768 GGGGTCTGTGGGGCCCTGTGGGCAAGATGGCATCCCTGGGAGTCTACACCTATATTTGCAA 827
Db 10924 GGGGTCTGTGGGGCCCTGTGGGCAAGATGGCATCCCTGGGAGTCTACACCTATATTTGCAA 10865
QY 828 GTATGTGGACTGGATCGGGATGATCATGAGGACAACTGACCTGTTTCTCGACCTCCAC 887
Db 10864 GTATGTGGACTGGATCGGGATGATCATGAGGACAACTGACCTGTTTCTCGACCTCCAC 10805
QY 888 CCCACGCCCTTAACCTTGGGTACCCCTCTGGCCCTCGAGACCAATATCTCTCCATCAC 947
Db 10804 CCCACGCCCTTAACCTTGGGTACCCCTCTGGCCCTCGAGACCAATATCTCTCCATCAC 10746
QY 948 TTCCCTTAGCTCCACTCTTGTGGCTGGGAACTTTTGGAACTTTAACTCTCGCCAGCC 1007
Db 10745 TTCCCTTAGCTCCACTCTTGTGGCTGGGAACTTTTGGAACTTTAACTCTCGCCAGCC 10686
QY 1008 CTTCTAAGACCCACGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATG 1067
Db 10685 CTTCTAAGACCCACGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATG 10626
QY 1068 AAGGAGGGGC 1077
Db 10625 AAGGAGGGGC 10616

RESULT 12
AX070938
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1. Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
      Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
      Lamson, G., Drmanac, R., Orkenjakov, R., Drmanac, S., Dickson, M.,
      Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.
      Human genes and gene expression products
      Patent: WO 0102568-A 1410 11-JAN-2001;
      CHIRON CORPORATION (US); HYSEQ, INC. (US)
  FEATURES
    source
      1. .405
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN
Query Match      20.8%; Score 227; DB 6; Length 405;
Best Local Similarity 100.0%; Pred. No. 6.3e-123;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 TTTTCCCCCAGACTTTGGAAAGTGACCCACCTGGGGCTCAGCATCTTTTGTCTGTGTGT 150
|||||
```

Db 91 TTTTCCCCCAGACTTTGGAAGTGACCCACCATGGGGCTCAGCATCTTTTGGCTCCTGTGT 150
 Qy 151 GTTCTTGGGCTCAGCCAGGAGCCACACCGAAGATTTTCAATGSCACTGAGTGTGGCGGT 210
 Db 151 GTTCTTGGGCTCAGCCAGGAGCCACACCGAAGATTTTCAATGSCACTGAGTGTGGCGGT 210
 Qy 211 AACTCAGCGGTGGCAGGTGGGCTCTTTGAGGGCACCGACCTGCGCTCGGGGGTGTTC 270
 Db 211 AACTCAGCGGTGGCAGGTGGGCTCTTTGAGGGCACCGACCTGCGCTCGGGGGTGTTC 270
 Qy 271 CTATTACACACAGGTGGTCTCTCACAGCGGCTCAGTCGACGGCGCAG 317
 Db 271 CTATTACACACAGGTGGTCTCTCACAGCGGCTCAGTCGACGGCGCAG 317

RESULT 13
 AC140096/c
 LOCUS 174724 bp DNA linear HTG 14-JUN-2003
 DEFINITION Pan troglodytes clone RP43-63K10, WORKING DRAFT SEQUENCE, 10
 ordered pieces.
 AC140096
 AC140096.4 GI:31745211
 VERSION HTG; HTGS PHASE2; HTGS DRAFT.
 KEYWORDS Pan troglodytes (chimpanzee)
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 Antonellis, A., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,
 Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
 Gupta, J., Haghighi, P., Harlan, J., Hansen, N., Ho, S.-L., Hu, P.,
 Hurlb, B., Iqbal, R., Karlins, P., Kwong, P., Latic, P., Lee-Fin, S.-Q.,
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C.,
 Maskeri, B., McDowell, J., Pagnirigan, C., Pearson, R., Portnoy, M.E.,
 Prasad, A., Reddix-Bugue, N., Schandler, K., Schueler, M.G., Shah, K.,
 Sison, C., Stantrypop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
 Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative

TITLE
 JOURNAL
 REFERENCE
 2 (bases 1 to 174724)
 Green, E.D.
 Direct Submission
 Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
 JOURNAL
 Submitted Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 174724)
 Green, E.D.
 Direct Submission
 Submitted (14-JUN-2003) NIH Intramural Sequencing Center, 8717
 JOURNAL
 Submitted Circle, Gaithersburg, MD 20877, USA
 On Jun 14, 2003 this sequence version replaced gi:29725783.

COMMENT
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nih.gov
 ----- Project Information
 Center project name: dxz
 Center clone name: 063K10

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 172841 bases at least Q40
 Consensus quality: 173490 bases at least Q30
 Consensus quality: 173737 bases at least Q20
 Insert size: 150000; agarose-Ep
 Insert size: 173824; sum-of-contigs
 Quality coverage: 15.91x in Q20 bases; agarose-fp
 Quality coverage: 13.73x in Q20 bases; sum-of-contigs

 * NONE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1
 47271 47270: contig of 47270 bp in length
 * 47371 58115: contig of 10745 bp in length
 * 58116 61050: contig of 2835 bp in length
 * 61051 65346: contig of 4196 bp in length
 * 65347 71536: contig of 6090 bp in length
 * 71537 83139: contig of 11503 bp in length
 * 83140 85133: contig of 1894 bp in length
 * 85134 88937: contig of 3604 bp in length
 * 88938 129108: contig of 40171 bp in length
 * 129109 129208: contig of 45516 bp in length.
 * 129209 174724: contig of 45516 bp in length.

FEATURES
 Location/Qualifiers
 1..174724
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone_lib="RP43-63K10"
 /clone_lib="RP43"
 1..79098
 /note="clone overlaps with GenBank Accession Number
 AC130782 clone CH251-355A20 (center project name dhz)"
 misc_feature
 1..47270
 /note="assembly_fragment
 clone_end:17
 vector_side:left"
 misc_feature
 47371..58115
 /note="assembly_fragment"
 misc_feature
 58216..61050
 /note="assembly_fragment"
 misc_feature
 61151..65346
 /note="assembly_fragment"
 misc_feature
 65447..71536
 /note="assembly_fragment"
 misc_feature
 71637..83139
 /note="assembly_fragment"
 misc_feature
 83240..85133
 /note="assembly_fragment"
 misc_feature
 85234..88837
 /note="assembly_fragment"
 misc_feature
 88938..129108
 /note="assembly_fragment"
 misc_feature
 113257..174724
 /note="clone overlaps with GenBank Accession Number
 AC133392 clone CH251-126O24 (center project name dxa)"
 misc_feature
 129209..174724

On Nov 22, 2002 this sequence version replaced gi:22218452.
----- Genome Center
Center: NIH Intramural Sequencing Center
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@hgr.nih.gov
----- Project Information
Center project name: dhz
Center clone name: 355A20

COMMENT

/note="assembly_fragment
clone_end:SP6
vector_side:right"

Query Match 19.9%; Score 217; DB 2; Length 174724;
Best Local Similarity 99.2%; Pred. No. 6.5e-117;
Matches 367; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 708 CCAGGTTGATCTGGGGCCCCCTGGTGTGGGGAGTCCCTCAAGTCTGGTCTCTG 767
DB 38225 CCAGGTTGATCTGGGGCCCCCTGGTGTGGGGAGTCCCTCAAGTCTGGTCTCTG 38166

QY 768 GGGTCTGTGGGGCCCTGGGACAGATGGCATCCCTGGAGTCTACACTATATTGCA 827
DB 38165 GGGTCTGTGGGGCCCTGGGACAGATGGCATCCCTGGAGTCTACACTATATTGCA 38106

QY 828 GTATGTGGACTGGATCCGGATCATGATGAGGAACAACTGACCTGTTTCTCCACCTCCAC 887
DB 38105 GTATGTGGACTGGATCCGGATCATGATGAGGAACAACTGACCTGTTTCTCCACCTCCAC 38046

QY 898 CCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 947
DB 38045 CCCACCCCTTAACCTGGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 37986

QY 948 TTCCTCTAGCTCCACTCTTTTGGCTGGAACTTCTTGGAACTTTAACTCTGCCAGCC 1007
DB 37985 TTCCTCTAGCTCCACTCTTTTGGCTGGAACTTCTTGGAACTTTAACTCTGCCAGCC 37926

QY 1008 CTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATAAATG 1067
DB 37925 CTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATAAATG 37866

QY 1068 AAGGAGGGGC 1077
DB 37865 AAGGAGGGGC 37856

RESULT 14
AC130782/c
LOCUS
DEFINITION Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12 ordered pieces.
AC130782
VERSION AC130782.2 GI:25167101
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1 (bases 1 to 200792)
AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Larc, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B., Marquis, E.H., Masello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A., and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 200792)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 200792)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 19656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 28306: contig of 28306 bp in length
* 28307 28406: gap of unknown length
* 28407 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length
* 37957 73522: contig of 35566 bp in length
* 73523 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 88817: contig of 5150 bp in length
* 88818 88917: gap of unknown length
* 88918 125611: contig of 36694 bp in length
* 125612 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174799: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193344: contig of 6862 bp in length
* 193345 193444: gap of unknown length
* 193445 199363: contig of 5919 bp in length
* 199364 199464: gap of unknown length
* 199464 200792: contig of 1329 bp in length.

Location/Qualifiers

1. .200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"

FEATURES
source

misc_feature 1. 28306
/note="assembly_fragment
clone_end:8p6
vector_side:left"
misc_feature 28407. 37856
/note="assembly_fragment"
misc_feature 37957. 73522
/note="assembly_fragment"
misc_feature 73623. 83567
/note="assembly_fragment"
misc_feature 83668. 88817
/note="assembly_fragment"
misc_feature 88918. 125611
/note="assembly_fragment"
misc_feature 125712. 159879
/note="assembly_fragment"
misc_feature 159980. 174698
/note="assembly_fragment"
misc_feature 174799. 186382
/note="assembly_fragment"
misc_feature 186483. 193344
/note="assembly_fragment"
misc_feature 193445. 199363
/note="assembly_fragment"
misc_feature 199464. 200792
/note="assembly_fragment"
clone_end:r7
vector_side:right"

ORIGIN

Query Match 19.9%; Score 217; DB 2; Length 200792;
Best Local Similarity 99.2%; Pred. No. 6.5e-117;
Matches 367; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 708 CCAGGTGATTCGGGGCCCCCTGGTGTGGGGAGTCCTTCAAGGTCTGGTCTCTG 767
Db 148657 CCAGGTGATTCGGGGCCCCCTGGTGTGGGGAGTCCTTCAAGGTCTGGTCTCTG 148598
QY 768 GGGGTCTGGGGCCCTGGGACAGATGGCATCCTGGAGTACACCTATATTGGAA 827
Db 148597 GGGGTCTGGGGCCCTGGGACAGATGGCATCCTGGAGTACACCTATATTGGAA 148538
QY 828 GTATGTGACTGGATCCGGATGATCATGAGGAACATGACCTGTTCTCCACCTCCAC 887
Db 148537 GTATGTGACTGGATCCGGATGATCATGAGGAACATGACCTGTTCTCCACCTCCAC 148478
QY 888 CCCACCCCTTAACTTGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 947
Db 148477 CCCACCCCTTAACTTGGTACCCCTCTGGCCCTCAGACACCAATATCTCTCCATCAC 148418
QY 948 TTCCCTAGCTCCACTCTGTGTGGCTGGAACTTCTTGGAACTTAACTCTCTGCCAGCC 1007
Db 148417 TTCCCTAGCTCCACTCTGTGTGGCTGGAACTTCTTGGAACTTAACTCTCTGCCAGCC 148358
QY 1008 CTTCTAAGACCCACGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATG 1067
Db 148357 CTTCTAAGACCCACGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATG 148298
QY 1068 AAGGAGGGGC 1077
Db 148297 AAGGAGGGGC 148288

RESULT 15
AC073185
LOCUS AC073185 142334 bp DNA linear HTG 17-AUG-2000
DEFINITION Homo sapiens chromosome 19 clone RP11-10111, WORKING DRAFT
SEQUENCE, 32 unordered pieces.
AC073185
ACCESSION AC073185.4 GI:9838034
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 142334)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 142334)
Waterston,R.H.
Direct Submission
Submitted (10-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 17, 2000 this sequence version replaced gi:9653152.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0010111
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-primer; 100%
Chemistry: Dye-terminator; Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: bases at least Q40
Consensus quality: bases at least Q30
Insert size: 135000; agarose-fp
Insert size: 139234; sum-of-contigs
Quality coverage: 3.86 in Q20 bases; agarose-fp
Quality coverage: 3.87 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1452: contig of 1452 bp in length
* 1453 1552: gap of unknown length
* 1553 2737: contig of 1185 bp in length
* 2738 2837: gap of unknown length
* 2838 4665: contig of 1828 bp in length
* 4666 4765: gap of unknown length
* 4766 6054: contig of 1289 bp in length
* 6055 6154: gap of unknown length
* 6155 7791: contig of 1637 bp in length
* 7792 7891: gap of unknown length
* 7892 10026: contig of 2135 bp in length
* 10027 10126: gap of unknown length
* 10127 11895: contig of 1769 bp in length
* 11896 11995: gap of unknown length
* 11996 15070: contig of 3075 bp in length
* 15071 15170: gap of unknown length
* 15171 18476: contig of 3106 bp in length
* 18477 18576: gap of unknown length
* 18577 21537: contig of 2961 bp in length
* 21538 21637: gap of unknown length
* 21638 24625: contig of 2388 bp in length
* 24626 24725: gap of unknown length
* 24726 27583: contig of 2858 bp in length
* 27584 27684: gap of unknown length
* 27684 30892: contig of 3209 bp in length
* 30893 35209: gap of unknown length
* 35210 35309: contig of 4217 bp in length
* 35310 39261: gap of unknown length
* 39261 39261: contig of 3952 bp in length
* 39262 42356: gap of unknown length
* 42356 42356: contig of 2995 bp in length

```
* 42357 42456: gap of unknown length
* 42457 46421: contig of 3965 bp in length
* 46422 46521: gap of unknown length
* 46522 51157: contig of 4636 bp in length
* 51158 51257: gap of unknown length
* 51258 57416: contig of 6159 bp in length
* 57417 57516: gap of unknown length
* 57517 61710: contig of 4194 bp in length
* 61711 61810: gap of unknown length
* 61811 66249: contig of 4439 bp in length
* 66250 66350: gap of unknown length
* 66351 71980: contig of 5631 bp in length
* 71981 72080: gap of unknown length
* 72081 77454: contig of 5374 bp in length
* 77455 77554: gap of unknown length
* 77555 84318: contig of 6764 bp in length
* 84319 84419: gap of unknown length
* 84420 88649: contig of 4231 bp in length
* 88650 88749: gap of unknown length
* 88750 94419: contig of 5670 bp in length
* 94420 100165: gap of unknown length
* 100166 100265: contig of 5646 bp in length
* 100266 107334: contig of 7069 bp in length
* 107335 107434: gap of unknown length
* 107435 115093: contig of 7659 bp in length
* 115094 115194: gap of unknown length
* 115195 122023: contig of 6830 bp in length
* 122024 122123: gap of unknown length
* 122124 132649: contig of 10526 bp in length
* 132650 132750: gap of unknown length
* 132751 142334: contig of 9585 bp in length.
```

FEATURES

source

```
1..142334
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="19"
  /clone="RP11-10111"
1..1452
  /note="assembly_name:Contig26"
1553..2737
  /note="assembly_name:Contig27"
2838..4665
  /note="assembly_name:Contig28"
4766..6054
  /note="assembly_name:Contig29"
6155..7791
  /note="assembly_name:Contig30"
7892..10026
  /note="assembly_name:Contig31"
10127..11895
  /note="assembly_name:Contig33"
11996..15070
  /note="assembly_name:Contig34"
15171..18476
  /note="assembly_name:Contig35"
18577..21537
  /note="assembly_name:Contig36"
21638..24625
  /note="assembly_name:Contig37"
24726..27583
  /note="assembly_name:Contig38"
27684..30892
  /note="assembly_name:Contig39"
30993..35209
  /note="assembly_name:Contig40"
35310..39261
  /note="assembly_name:Contig41"
39362..42356
  /note="assembly_name:Contig42"
42457..46421
  /note="assembly_name:Contig43"
```

```
misc_feature 46522..51157
              /note="assembly_name:Contig44"
misc_feature 51258..57416
              /note="assembly_name:Contig45"
misc_feature 57517..61710
              /note="assembly_name:Contig46"
misc_feature 61811..66249
              /note="assembly_name:Contig47"
misc_feature 66350..71980
              /note="assembly_name:Contig48"
misc_feature 72081..77454
              /note="assembly_name:Contig49"
misc_feature 77555..84318
              /note="assembly_name:Contig50"
misc_feature 84419..88649
              /note="assembly_name:Contig51"
misc_feature 88750..94419
              /note="assembly_name:Contig52"
misc_feature 94520..100165
              /note="assembly_name:Contig53"
misc_feature 100266..107334
              /note="assembly_name:Contig54"
misc_feature 107435..115093
              /note="assembly_name:Contig55"
misc_feature 115194..122023
              /note="assembly_name:Contig56"
misc_feature 122124..132649
              /note="assembly_name:Contig57"
              clone_end:T7
              vector_side:right
misc_feature 132750..142334
              /note="assembly_name:Contig58"
```

ORIGIN

```
Query Match      11.6%; Score 127; DB 2; Length 142334;
Best Local Similarity 100.0%; Pred.No. 3.1e-63;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 578 ACCCATCCCGGATCTGCTCCAGTGCCTCAACCTTCATCGTCTCCCATGCCACCTGCC 637
    |||
Db 3965 ACCCATCCCGGATCTGCTCCAGTGCCTCAACCTTCATCGTCTCCCATGCCACCTGCC 4024

QY 638 ATGGTGTGTATCCCGGGAGAAATCAGAGCAACATGGTGTGTGCGAGCGGGCGTCCCGGGGC 697
    |||
Db 4025 ATGGTGTGTATCCCGGGAGAAATCAGAGCAACATGGTGTGTGCGAGCGGGCGTCCCGGGGC 4084

QY 698 AGGATGC 704
    |||
Db 4085 AGGATGC 4091
```

Search completed: June 20, 2004, 07:34:57
Job time : 4629 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 04:42:21 ; Search time 509 Seconds
(without alignments)
9105.666 Million cell updates/sec

Title: US-10-015-385A-193
Perfect score: 1091
Sequence: 1 caagcaggtcatcccttg.....aggggcaaaaaaaaaa 1091

Scoring table: OLIGO NUC
Gapop_60.0, Gapext 60.0

Searched: 3373863 seqs, 212409041 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1091	100.0	1091	3	AAA77671 Human PRO
2	1091	100.0	1091	3	AAC58114 Human PRO
3	1091	100.0	1091	3	AAA37075 Human PRO
4	1091	100.0	1091	4	AAF54341 DNA encod
5	1091	100.0	1091	8	ACD68378 Novel hum
6	1091	100.0	1091	8	ACH04480 Human CDN
7	1091	100.0	1091	8	ACD68024 Novel hum
8	1091	100.0	1091	9	ADC18062 Human PRO
9	1091	100.0	1091	9	ADD70708 Human CDN
10	1091	100.0	1091	9	ADD39785 Human CDN
11	1091	100.0	1091	9	ADD70231 Human CDN
12	1091	100.0	1091	9	ADD38352 Human CDN
13	1091	100.0	1091	9	ADD39308 Human CDN
14	1091	100.0	1091	9	ADD38831 Human CDN
15	1091	100.0	1091	9	ADD40262 Human CDN
16	1091	100.0	1091	9	AD50483 Human CDN
17	1091	100.0	1091	9	AD520095 Human CDN
18	1091	100.0	1091	9	AD550006 Human CDN
19	1091	100.0	1091	9	AD521564 Human CDN
20	963	88.3	1185	4	AAH98653 Human EST
21	727	66.6	945	6	ABK94900 Human nov
22	696	63.8	747	7	AAL59914 Human kal
23	487	44.6	619	6	ABK30369 Human G-p

ALIGNMENTS

RESULT 1

AAA77671	440	40.3	840	2	AAAX80906	11820	3	AAAX80906	Human CDN
XX	AAA77671	standard; cDNA; 1091 BP.			AAAX80906	11820	3	AAAX80906	Human CDN
AC	AAA77671				AAAX80906	11820	3	AAAX80906	Human CDN
XX	AAA77671				AAAX80906	11820	3	AAAX80906	Human CDN
DT	07-NOV-2000	(first entry)			AAAX80906	11820	3	AAAX80906	Human CDN
XX	Human PRO1303	cDNA sequence SEQ ID NO:202.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;				AAAX80906	11820	3	AAAX80906	Human CDN
KW	diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;				AAAX80906	11820	3	AAAX80906	Human CDN
KW	angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;				AAAX80906	11820	3	AAAX80906	Human CDN
KW	cytostatic; gene therapy; vaccine; ss.				AAAX80906	11820	3	AAAX80906	Human CDN
XX	Homo sapiens.				AAAX80906	11820	3	AAAX80906	Human CDN
OS	Homo sapiens.				AAAX80906	11820	3	AAAX80906	Human CDN
XX	WO2000032221-A2.				AAAX80906	11820	3	AAAX80906	Human CDN
XX	08-JUN-2000.				AAAX80906	11820	3	AAAX80906	Human CDN
XX	30-NOV-1999;	99WO-US028313.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	01-DEC-1998;	98WO-US025108.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	16-DEC-1998;	98US-0112850P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	12-JAN-1999;	99US-0115554P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	08-MAR-1999;	99WO-US005028.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	12-MAR-1999;	99US-0123957P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	28-APR-1999;	99US-0131445P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	14-MAY-1999;	99US-0134287P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	02-JUN-1999;	99WO-US012252.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	23-JUN-1999;	99US-0141037P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	20-JUL-1999;	99US-0144758P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	26-JUL-1999;	99US-0145698P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	01-SEP-1999;	99WO-US020111.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	08-SEP-1999;	99WO-US020594.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	13-SEP-1999;	99WO-US020944.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	15-SEP-1999;	99WO-US021090.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	15-SEP-1999;	99WO-US021547.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	05-OCT-1999;	99WO-US023089.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	29-OCT-1999;	99US-0162506P.			AAAX80906	11820	3	AAAX80906	Human CDN
XX	(GETH) GENENTECH INC.				AAAX80906	11820	3	AAAX80906	Human CDN
XX	Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ;				AAAX80906	11820	3	AAAX80906	Human CDN
XX	Goddard A, Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF;				AAAX80906	11820	3	AAAX80906	Human CDN

XX Claim 50; Fig 23; 226pp; English.

XX

XX The present invention describes an antibody that binds to a human protein

CC (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780; PRO3434;

CC PRO1927; PRO3567; PRO1293; PRO1303; PRO4344; PRO4354; PRO4397;

CC PRO4407; PRO1555; PRO1096; PRO2038; and PRO2262. (I) has anticancer

CC activity and can be used to diagnose tumours in mammals, by detecting

CC complex formation when the antibody is contacted with test cells.

CC increased expression of genes encoding (I) can also be detected to

CC diagnose tumours. Agents which inhibit the activity of (I), especially

CC the antibodies, or an antisense oligonucleotide which hybridises to genes

CC encoding (I), can be used to inhibit tumour growth, preferably by

CC inducing cell death. Methods from the present invention can be used to

CC identify compounds which inhibit the biological activity of (I). AAC58019

CC identifies representative PCR primers and hybridisation probes used in

CC examples from the present invention for human PRO sequences. AAC58103 to

CC AAC58122 and AAB24021 to AAB24040 represent human PRO polynucleotide and

CC protein sequences given in the exemplification of the present invention

XX

XX Sequence 1091 BP; 221 A; 335 C; 222 T; 0 U; 0 Other;

XX

Qy	721	GGGGGCCCTCTGTGTGTGGGGAGTCCCTTCAAGGTCTGGTGTCTGGGGGTCTGTGGGG	780
Db	721	GGGGGCCCTCTGTGTGTGGGGAGTCCCTTCAAGGTCTGGTGTCTGGGGGTCTGTGGGG	780
Qy	781	CCCTGTGGACAAGATGGCATCCCTCGAGTCTACACCTATATTTGCAAGTATGTGGACTGG	840
Db	781	CCCTGTGGACAAGATGGCATCCCTCGAGTCTACACCTATATTTGCAAGTATGTGGACTGG	840
Qy	841	ATCCGATGATCATGAGGAACAACACTGACCTGTTTCTCCACCTCCACCCCCACCCCTTAA	900
Db	841	ATCCGATGATCATGAGGAACAACACTGACCTGTTTCTCCACCTCCACCCCCACCCCTTAA	900
Qy	901	CTTGGGTACCCCTCTGGCCCTCAGAGACAATATCTCTCATCTTCCCTAGTCTCCCTAGTCTC	960
Db	901	CTTGGGTACCCCTCTGGCCCTCAGAGACAATATCTCTCATCTTCCCTAGTCTCCCTAGTCTC	960
Qy	961	ACTCTTCTTGGCCCTGGGGAACCTCTTGGAACTTTAACTCTGCCAGCCCTTCTTAAGACCCA	1020
Db	961	ACTCTTCTTGGCCCTGGGGAACCTCTTGGAACTTTAACTCTGCCAGCCCTTCTTAAGACCCA	1020
Qy	1021	CGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATATAAATGAAGAGGGGCAAA	1080
Db	1021	CGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGAATAAATATAAATGAAGAGGGGCAAA	1080
Qy	1081	AAAAAAAAAAAA 1091	
Db	1081	AAAAAAAAAAAA 1091	
RESULT 3			
AAA37075			
ID	AAA37075 standard; cDNA; 1091 BP.		
XX	AAA37075;		
XX	XX		
DT	08-AUG-2000 (first entry)		
XX	XX		
DE	Human PRO1303 (UNQ669) cDNA sequence SEQ ID NO:193.		
XX	XX		
KW	Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;		
XX	transmembrane; secretion; immunoadhesion; pharmaceutical; screening; ss.		
OS	Homo sapiens.		
XX	XX		
PN	WO200012708-A2.		
XX	XX		
PD	09-MAR-2000.		
XX	XX		
PF	01-SEP-1999; 99WO-US020111.		
XX	XX		
PR	01-SEP-1998; 98US-0098716P.		
PR	01-SEP-1998; 98US-0098749P.		
PR	01-SEP-1998; 98US-0098750P.		
PR	02-SEP-1998; 98US-0098803P.		
PR	02-SEP-1998; 98US-0098821P.		
PR	02-SEP-1998; 98US-0098843P.		
PR	09-SEP-1998; 98US-0099536P.		
PR	09-SEP-1998; 98US-0099596P.		
PR	09-SEP-1998; 98US-0099598P.		
PR	09-SEP-1998; 98US-0099602P.		
PR	09-SEP-1998; 98US-0099642P.		
PR	10-SEP-1998; 98US-0099741P.		
PR	10-SEP-1998; 98US-0099754P.		
PR	10-SEP-1998; 98US-0099763P.		
PR	10-SEP-1998; 98US-0099792P.		
PR	10-SEP-1998; 98US-0099808P.		
PR	10-SEP-1998; 98US-0099812P.		
PR	10-SEP-1998; 98US-0099815P.		
PR	10-SEP-1998; 98US-0099816P.		
PR	15-SEP-1998; 98US-0100385P.		
PR	15-SEP-1998; 98US-0100388P.		
PR	15-SEP-1998; 98US-0100390P.		

```
PR 16-SEP-1998; 98US-0100584P.
PR 16-SEP-1998; 98US-0100627P.
PR 16-SEP-1998; 98US-0100661P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100663P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100710P.
PR 17-SEP-1998; 98US-0100711P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100848P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 22-SEP-1998; 98US-0101071P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101473P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101476P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 01-OCT-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 02-OCT-1998; 98US-0102687P.
PR 06-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 07-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 08-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 14-OCT-1998; 98US-0103711P.
PR 20-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 21-OCT-1998; 98US-0105002P.
PR 22-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 26-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105861P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.

PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 10-NOV-1998; 98US-0106934P.
PR 17-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
XX
XX
XX
PI Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI; 2000-237871/20.
XX P-PSDB; AAY99393.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or secreted
PT PRO polypeptides, useful for screening of potential peptide or small
PT molecule inhibitors of the relevant receptor/ligand interactions.
XX
PS Claim 2; Fig 107; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding then have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR
CC primers and hybridisation probes used in the isolation of the PRO
XX polypeptides from the present invention
SQ Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 1091; DB 3; Length 1091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAGCAGGTATCCCTTGGTGTACCTTCAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGC 60
Db |||||
Db 1 CAAGCAGGTATCCCTTGGTGTACCTTCAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGC 60
Qy 61 ACAGGGAAGAGGTGACCTCTGAGATTCCTTCCCTTCCCTCCAGACTTTGGAAGTACCCACC 120
Db |||||
Db 61 ACAGGGAAGAGGTGACCTCTGAGATTCCTTCCCTTCCCTCCAGACTTTGGAAGTACCCACC 120
Qy 121 ATGGGGCTCAGCATCTTTTGTCTTGTGTTCTTTGGGCTCAGCAGGCGAGCCACCG 180
Db |||||
Db 121 ATGGGGCTCAGCATCTTTTGTCTTGTGTTCTTTGGGCTCAGCAGGCGAGCCACCG 180
Qy 181 AAGATTTTCAATGGGCTAGTGTTGGCGGTAACTCAAGCCGTGGCAGGTGGGGCTGTTT 240
Db |||||
Db 181 AAGATTTTCAATGGGCTAGTGTTGGCGGTAACTCAAGCCGTGGCAGGTGGGGCTGTTT 240
```



```
QY 241 GAGGACACAGCTGCGTCCGGGGTGTCTTATTGACACAGGTGGTCTCTACAGCG 300
Db 241 GAGGGACACAGCTGCGTCCGGGGTGTCTTATTGACACAGGTGGTCTCTACAGCG 300
QY 301 GCTCACTGACGGCAGCAGGTAAGTCTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
Db 301 GCTCACTGACGGCAGCAGGTAAGTCTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
QY 361 GACTGACCGGACAGATCCGGACACAGCGCTCTCTGTGACCCATCCCGGCTACCTGGGA 420
Db 361 GACTGACCGGACAGATCCGGACACAGCGCTCTCTGTGACCCATCCCGGCTACCTGGGA 420
QY 421 GCCTGACGAGCCAGCAGCAGCAGCTCCGGCTGCTGGGCTGGGCTGCGCTGCGCTGCGCTA 480
Db 421 GCCTGACGAGCCAGCAGCAGCAGCTCCGGCTGCTGGGCTGGGCTGCGCTGCGCTGCGCTA 480
QY 481 ACCAGCAGGTTCAACCCCTGCCCTGCCCTGCGCTGCTGCAACCGCTGGCACCGAGTGC 540
Db 481 ACCAGCAGGTTCAACCCCTGCCCTGCCCTGCGCTGCTGCAACCGCTGGCACCGAGTGC 540
QY 541 CACGTCTCAGCTGGGCGATCACCAACACACACCGAACCCATTCCTCGGATCTGCTCCAG 600
Db 541 CACGTCTCAGCTGGGCGATCACCAACACACACCGAACCCATTCCTCGGATCTGCTCCAG 600
QY 601 TGCTCAACCTCTCCATGCTCTCCATGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 TGCTCAACCTCTCCATGCTCTCCATGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 ACAGGACACATGTTGTGAGGCGGCTGCTGCGGCGAGGATGCTGCGAGGCTGATCTCT 720
Db 661 ACAGGACACATGTTGTGAGGCGGCTGCTGCGGCGAGGATGCTGCGAGGCTGATCTCT 720
QY 721 GGGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 GGGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 CCTGTGGAACAAGTGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGAGTCTG 840
Db 781 CCTGTGGAACAAGTGGCATCCCTGGAGTCTACACCTATATTGCAAGTATGTGAGTCTG 840
QY 841 ATCGGATGATCATGAGAACATGACCTGTTTCTCCACCTCCACCCCCACCCCTTAA 900
Db 841 ATCGGATGATCATGAGAACATGACCTGTTTCTCCACCTCCACCCCCACCCCTTAA 900
QY 901 CTTGGGTACCCCTCTGCGCCCTCAGACACCAATATCTCTCCATCACTTCCCTAGCTCC 960
Db 901 CTTGGGTACCCCTCTGCGCCCTCAGACACCAATATCTCTCCATCACTTCCCTAGCTCC 960
QY 961 ACTCTGTTGGCTGGGAACTTCTTGGAACCTTTAACTCTCCAGCCCTTCTTAAGACCCA 1020
Db 961 ACTCTGTTGGCTGGGAACTTCTTGGAACCTTTAACTCTCCAGCCCTTCTTAAGACCCA 1020
QY 1021 CGAGCGGGGTGAGAGAGTGTCAATAGTCTGGATTAATATAATGAAGAGGGGCAAA 1080
Db 1021 CGAGCGGGGTGAGAGAGTGTCAATAGTCTGGATTAATATAATGAAGAGGGGCAAA 1080
QY 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091
```

RESULT 4

AAF54341
ID AAF54341 standard; DNA; 1091 BP.

XX AAF54341;

AC AAF54341;

DT 02-APR-2001 (first entry)

XX DNA encoding protein of the invention #54.

DE Secreted; transmembrane; gene therapy; ss.

XX

Unidentified.
WO200078961-A1.
28-DEC-2000.
18-FEB-2000; 2000WO-US004342.
23-JUN-1999; 99US-0141037P.
20-JUL-1999; 99US-0144758P.
26-JUL-1999; 99US-0145698P.
01-SEP-1999; 99WO-US020111.
29-OCT-1999; 99US-0162506P.
30-NOV-1999; 99WO-US028313.
02-DEC-1999; 99WO-US028551.
16-DEC-1999; 99WO-US030095.
05-JAN-2000; 2000WO-US000219.
06-JAN-2000; 2000WO-US000376.
(GETH) GENENTECH INC.
Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
Williams PM, Wood WI;
WPI; 2001-071395/08.

Secreted and transmembrane proteins and nucleic acids designated PRO,
useful as hybridization probes, in chromosome and gene mapping and gene
therapy.

Claim 2; Fig 107; 787pp; English.

The present invention relates to secreted and transmembrane proteins.
These proteins and the DNA encoding them may be used as hybridization
probes, in chromosome and gene mapping and in the generation of anti-
sense RNA and DNA. They may also be used to generate either of anti-
transgenic animals or knockout animals which are in turn useful for
development and screening of therapeutically useful reagents. The nucleic
acids may also be used in gene therapy

Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;

Query Match 100.0%; Score 1091; DB 4; Length 1091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAAGCAGGTCAATCCCTTGGTGAACCTTCAAGAGAAAGCAGAGAGGCGGAGGCGGCGG 60
Db 1 CAAGCAGGTCAATCCCTTGGTGAACCTTCAAGAGAAAGCAGAGAGGCGGAGGCGGCGG 60
QY 61 ACAGGGAAGGGTGACCTCTGAGATTCCTTTCCCTCCAGAGCTTTGGAAGTCAACCCACC 120
Db 61 ACAGGGAAGGGTGACCTCTGAGATTCCTTTCCCTCCAGAGCTTTGGAAGTCAACCCACC 120
QY 121 ATGGGGCTCAGCATCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 ATGGGGCTCAGCATCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAACTCAACAGCCGTGGCGGTGTTT 240
Db 181 AAGATTTTCAATGGCACTGAGTGTGGGCGTAACTCAACAGCCGTGGCGGTGTTT 240
QY 241 GAGGCGCAGCCTGCGCTGCGGGGTGTCTTATGACCAAGGTGGGTCTCTACAGCG 300
Db 241 GAGGCGCAGCCTGCGCTGCGGGGTGTCTTATGACCAAGGTGGGTCTCTACAGCG 300
QY 301 GCTCACTGACGGCAGCAGGTAAGTCTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
Db 301 GCTCACTGACGGCAGCAGGTAAGTCTGGGTGCGCTGGGGGAAACACAGCCTCAGCCAGCTC 360
QY 361 GACTGACCGGACAGATCCGGACACAGCGCTCTCTGTGACCCATCCCGGCTACCTGGGA 420
```



```
Db 601 |||||TGCTCAACCTCTCCATGCTCTCCATGCCATGCCATGGTGTATCCCGGGAATC 560
Qy 661 ACAGCAACATGCTGTCAGCGCGGTCCTCCCGGCGAGATGCCCTGCAGGGTGATTC 720
Db 661 ACAGCAACATGCTGTCAGCGCGGTCCTCCCGGCGAGATGCCCTGCAGGGTGATTC 720
Qy 721 GGGGGCCCCCTGCTGTGTGGGGGAGTCCTTCAAGGTCCTGTCCTGGGGGCTGTGGGG 780
Db 721 GGGGGCCCCCTGCTGTGTGGGGGAGTCCTTCAAGGTCCTGTCCTGGGGGCTGTGGGG 780
Qy 781 CCTGTGGAAGATGCGCATCCCTGGAGTCTACACCTATATTGGCAAGTATGTGACTGG 840
Db 781 CCTGTGGAAGATGCGCATCCCTGGAGTCTACACCTATATTGGCAAGTATGTGACTGG 840
Qy 841 ATCCGGATCATCATGAGGAACCACTGACCTGTTCTCCACCTCCACCCCGCCCTTAA 900
Db 841 ATCCGGATCATCATGAGGAACCACTGACCTGTTCTCCACCTCCACCCCGCCCTTAA 900
Qy 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACAATATCTCTCCATCACTTCCCTAGCTCC 960
Db 901 CTTGGGTACCCCTCTGGCCCTCAGAGCAACAATATCTCTCCATCACTTCCCTAGCTCC 960
Qy 961 ACTCTGTGTGGCTGGGACTTCTTGGAACTTAACTCTGCGAGCCCTTCTAAGACCA 1020
Db 961 ACTCTGTGTGGCTGGGACTTCTTGGAACTTAACTCTGCGAGCCCTTCTAAGACCA 1020
Qy 1021 CGAGCGGGGTGAGAGAGTGTCAATAGTCTGGAATAAATAAATGAAGGAGGGGCAAA 1080
Db 1021 CGAGCGGGGTGAGAGAGTGTCAATAGTCTGGAATAAATAAATGAAGGAGGGGCAAA 1080
Qy 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091

RESULT 7
ACD68024
ID ACD68024 standard; cDNA; 1091 BP.
XX AC ACD68024;
XX DT 17-SEP-2003 (first entry)
XX DE Novel human secreted and transmembrane protein PRO1303 cDNA.
XX KW Human; secreted and transmembrane protein; PRO; gene therapy; vaccine;
XX KW tissue typing; chromosome identification; vaccine; gene; ss.
XX OS Homo sapiens.
XX FN US2003073129-A1.
XX PD 17-APR-2003.
XX PF 04-SEP-2001; 2001US-00946374.
XX PR 01-SEP-1998; 98US-0098716P.
XX PR 01-SEP-1998; 98US-0098723P.
XX PR 01-SEP-1998; 98US-0098749P.
XX PR 01-SEP-1998; 98US-0098750P.
XX PR 02-SEP-1998; 98US-0098803P.
XX PR 02-SEP-1998; 98US-0098821P.
XX PR 02-SEP-1998; 98US-0098843P.
XX PR 09-SEP-1998; 98US-0099536P.
XX PR 09-SEP-1998; 98US-0099596P.
XX PR 09-SEP-1998; 98US-0099598P.
XX PR 09-SEP-1998; 98US-0099602P.
XX PR 10-SEP-1998; 98US-0099642P.
XX PR 10-SEP-1998; 98US-0099741P.
XX PR 10-SEP-1998; 98US-0099754P.
XX PR 10-SEP-1998; 98US-0099763P.
XX PR 10-SEP-1998; 98US-0099792P.
PR 10-SEP-1998; 98US-0099808P.
PR 10-SEP-1998; 98US-0099812P.
PR 10-SEP-1998; 98US-0099815P.
PR 10-SEP-1998; 98US-0099816P.
PR 15-SEP-1998; 98US-0100385P.
PR 15-SEP-1998; 98US-0100388P.
PR 15-SEP-1998; 98US-0100390P.
PR 16-SEP-1998; 98US-0100584P.
PR 16-SEP-1998; 98US-0100627P.
PR 16-SEP-1998; 98US-0100661P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100710P.
PR 17-SEP-1998; 98US-0100711P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100848P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 18-SEP-1998; 98US-0101071P.
PR 22-SEP-1998; 98US-0101279P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101474P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101476P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101479P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 08-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
```



```
Db      721  GGGGGCCCCCTGGTGTGGGGAGTCTTCAAGTCTGGTCTCTGGGGTCTGTGGG 780
QY      781  CCTGTGGACAAGATGGCATCCTGGAGTCTACACCTATATTGCAAGTATGTGACTGG 840
Db      781  CCTGTGGACAAGATGGCATCCTGGAGTCTACACCTATATTGCAAGTATGTGACTGG 840
QY      841  ATCCGATGATCATGAGGAACTGACCTGTTCTCCACCTCCACCCCAACCCCTTAA 900
Db      841  ATCCGATGATCATGAGGAACTGACCTGTTCTCCACCTCCACCCCAACCCCTTAA 900
QY      901  CTTGGTACCCCTCTGGCCCTCAGAGCAACCAATCTCTCATCATCTCCCTAGTCC 960
Db      901  CTTGGTACCCCTCTGGCCCTCAGAGCAACCAATCTCTCATCATCTCCCTAGTCC 960
QY      961  ACTCTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTGCGAGCCCTTCTTAAGACCA 1020
Db      961  ACTCTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTGCGAGCCCTTCTTAAGACCA 1020
QY      1021  CGAGCGGGGTGAGAAAGTGTGCAATAGTCTGGAATAAATAAATGAAGGAGGGGCAA 1080
Db      1021  CGAGCGGGGTGAGAAAGTGTGCAATAGTCTGGAATAAATAAATGAAGGAGGGGCAA 1080
QY      1081  AAAAAAAAAA 1091
Db      1081  AAAAAAAAAA 1091

RESULT 8
ID      ADC18062
XX      ADC18062 standard; cDNA; 1091 BP.
AC      ADC18062;
XX
DT      18-DEC-2003 (first entry)
XX
DE      Human PRO polynucleotide #54.
KW      Human; PRO; gene; ss; protein electrophoresis; chromosome mapping;
KW      gene mapping; genetic disorder.
XX
OS      Homo sapiens.
XX
XX      US2003064925-A1.
XX
XX      03-APR-2003.
XX
XX      10-DEC-2001; 2001US-00013907.
XX
XX      01-SEP-1998; 98US-0098716P.
XX      01-SEP-1998; 98US-0098723P.
XX      01-SEP-1998; 98US-0098749P.
XX      02-SEP-1998; 98US-0098750P.
XX      02-SEP-1998; 98US-0098803P.
XX      02-SEP-1998; 98US-0098821P.
XX      02-SEP-1998; 98US-0098843P.
XX      03-SEP-1998; 98US-0099536P.
XX      03-SEP-1998; 98US-0099596P.
XX      03-SEP-1998; 98US-0099598P.
XX      03-SEP-1998; 98US-0099602P.
XX      03-SEP-1998; 98US-0099642P.
XX      10-SEP-1998; 98US-0099741P.
XX      10-SEP-1998; 98US-0099754P.
XX      10-SEP-1998; 98US-0099763P.
XX      10-SEP-1998; 98US-0099792P.
XX      10-SEP-1998; 98US-0099808P.
XX      10-SEP-1998; 98US-0099812P.
XX      10-SEP-1998; 98US-0099815P.
XX      10-SEP-1998; 98US-0099816P.
XX      15-SEP-1998; 98US-0100385P.
XX      15-SEP-1998; 98US-0100388P.
XX      15-SEP-1998; 98US-0100390P.
XX      16-SEP-1998; 98US-0100584P.
XX      16-SEP-1998; 98US-0100627P.
XX      16-SEP-1998; 98US-0100661P.
XX      16-SEP-1998; 98US-0100662P.
XX      16-SEP-1998; 98US-0100664P.
XX      17-SEP-1998; 98US-0100683P.
XX      17-SEP-1998; 98US-0100684P.
XX      17-SEP-1998; 98US-0100710P.
XX      17-SEP-1998; 98US-0100711P.
XX      17-SEP-1998; 98US-0100919P.
XX      17-SEP-1998; 98US-0100930P.
XX      18-SEP-1998; 98US-0100848P.
XX      18-SEP-1998; 98US-0100849P.
XX      18-SEP-1998; 98US-0101014P.
XX      18-SEP-1998; 98US-0101068P.
XX      18-SEP-1998; 98US-0101071P.
XX      22-SEP-1998; 98US-0101279P.
XX      23-SEP-1998; 98US-0101471P.
XX      23-SEP-1998; 98US-0101472P.
XX      23-SEP-1998; 98US-0101474P.
XX      23-SEP-1998; 98US-0101475P.
XX      23-SEP-1998; 98US-0101476P.
XX      23-SEP-1998; 98US-0101477P.
XX      24-SEP-1998; 98US-0101738P.
XX      24-SEP-1998; 98US-0101741P.
XX      24-SEP-1998; 98US-0101743P.
XX      24-SEP-1998; 98US-0101915P.
XX      24-SEP-1998; 98US-0101916P.
XX      29-SEP-1998; 98US-0102207P.
XX      29-SEP-1998; 98US-0102240P.
XX      29-SEP-1998; 98US-0102307P.
XX      29-SEP-1998; 98US-0102330P.
XX      30-SEP-1998; 98US-0102331P.
XX      30-SEP-1998; 98US-0102484P.
XX      30-SEP-1998; 98US-0102570P.
XX      30-SEP-1998; 98US-0102571P.
XX      01-OCT-1998; 98US-0102684P.
XX      01-OCT-1998; 98US-0102887P.
XX      02-OCT-1998; 98US-0102965P.
XX      06-OCT-1998; 98US-0103258P.
XX      06-OCT-1998; 98US-0103449P.
XX      07-OCT-1998; 98US-0103314P.
XX      07-OCT-1998; 98US-0103315P.
XX      07-OCT-1998; 98US-0103288P.
XX      07-OCT-1998; 98US-0103395P.
XX      07-OCT-1998; 98US-0103396P.
XX      08-OCT-1998; 98US-0103401P.
XX      08-OCT-1998; 98US-0103633P.
XX      08-OCT-1998; 98US-0103678P.
XX      08-OCT-1998; 98US-0103679P.
XX      14-OCT-1998; 98US-0103711P.
XX      20-OCT-1998; 98US-0104257P.
XX      20-OCT-1998; 98US-0104987P.
XX      20-OCT-1998; 98US-0105000P.
XX      21-OCT-1998; 98US-0105002P.
XX      21-OCT-1998; 98US-0105104P.
XX      22-OCT-1998; 98US-0105169P.
XX      22-OCT-1998; 98US-0105266P.
XX      26-OCT-1998; 98US-0105693P.
XX      26-OCT-1998; 98US-0105694P.
XX      27-OCT-1998; 98US-0105807P.
XX      27-OCT-1998; 98US-0105881P.
XX      27-OCT-1998; 98US-0105882P.
XX      28-OCT-1998; 98US-0106023P.
XX      28-OCT-1998; 98US-0106029P.
XX      28-OCT-1998; 98US-0106030P.
XX      28-OCT-1998; 98US-0106032P.
XX      28-OCT-1998; 98US-0106033P.
XX      28-OCT-1998; 98US-0106178P.
XX      29-OCT-1998; 98US-0106248P.
XX      29-OCT-1998; 98US-0106384P.
```



```
Db      841 ATCCGGATCATGATGAGGAACACTGACCTGTTTCTCCACCTCCACCCCTTAA 900
QY      901 CTTGGGTACCCCTCTGGCCCTCAGAGCACAATATCTCTCCATCACTTCCCTAGCTCC 960
Db      901 CTTGGGTACCCCTCTGGCCCTCAGAGCACAATATCTCTCCATCACTTCCCTAGCTCC 960
QY      961 ACTCTTGTGGCTGGGAACCTCTTGGAACTTTTAACCTCTGCGAGCCCTCTAAGACCCA 1020
Db      961 ACTCTTGTGGCTGGGAACCTCTTGGAACTTTTAACCTCTGCGAGCCCTCTAAGACCCA 1020
QY      1021 CGAGCGGGTGTGAGAGAGTGTCGAATAGTCTGGAATAAATAAATAAAGAGGGGCAAA 1080
Db      1021 CGAGCGGGTGTGAGAGAGTGTCGAATAGTCTGGAATAAATAAATAAAGAGGGGCAAA 1080
QY      1081 AAAAAAAAAA 1091
Db      1081 AAAAAAAAAA 1091

RESULT 9
ADD70708
ID ADD70708 standard; cDNA, 1091 bp.
AC ADD70708;
XX
XX
XX 15-JAN-2004 (first entry)
DE Human cDNA encoding secreted/transmembrane protein PRO1303.
KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX Homo sapiens.
XX
XX US200309625-A1.
XX
XX 29-MAY-2003.
XX
XX 12-DEC-2001; 2001US-00015386.
XX
XX 01-SEP-1998; 98US-0098716P.
XX 01-SEP-1998; 98US-0098723P.
XX 01-SEP-1998; 98US-0098749P.
XX 01-SEP-1998; 98US-0098750P.
XX 02-SEP-1998; 98US-0098803P.
XX 02-SEP-1998; 98US-0098821P.
XX 02-SEP-1998; 98US-0098843P.
XX 02-SEP-1998; 98US-00989536P.
XX 02-SEP-1998; 98US-00989596P.
XX 02-SEP-1998; 98US-00989598P.
XX 02-SEP-1998; 98US-00989602P.
XX 02-SEP-1998; 98US-00989642P.
XX 02-SEP-1998; 98US-00989741P.
XX 02-SEP-1998; 98US-00989754P.
XX 02-SEP-1998; 98US-00989763P.
XX 02-SEP-1998; 98US-00989792P.
XX 02-SEP-1998; 98US-00989808P.
XX 02-SEP-1998; 98US-00989812P.
XX 02-SEP-1998; 98US-00989815P.
XX 02-SEP-1998; 98US-00989816P.
XX 02-SEP-1998; 98US-0100385P.
XX 02-SEP-1998; 98US-0100388P.
XX 02-SEP-1998; 98US-0100390P.
XX 02-SEP-1998; 98US-0100384P.
XX 02-SEP-1998; 98US-0100627P.
XX 02-SEP-1998; 98US-0100661P.
XX 02-SEP-1998; 98US-0100662P.
XX 02-SEP-1998; 98US-0100664P.
XX
XX 01-SEP-1998; 98US-0100683P.
XX 01-SEP-1998; 98US-0100684P.
XX 01-SEP-1998; 98US-0100710P.
XX 01-SEP-1998; 98US-0100711P.
XX 01-SEP-1998; 98US-0100919P.
XX 01-SEP-1998; 98US-0100930P.
XX 01-SEP-1998; 98US-0100848P.
XX 01-SEP-1998; 98US-0100849P.
XX 01-SEP-1998; 98US-0101014P.
XX 01-SEP-1998; 98US-0101068P.
XX 01-SEP-1998; 98US-0101071P.
XX 01-SEP-1998; 98US-0101279P.
XX 01-SEP-1998; 98US-0101471P.
XX 01-SEP-1998; 98US-0101472P.
XX 01-SEP-1998; 98US-0101474P.
XX 01-SEP-1998; 98US-0101475P.
XX 01-SEP-1998; 98US-0101476P.
XX 01-SEP-1998; 98US-0101477P.
XX 01-SEP-1998; 98US-0101479P.
XX 01-SEP-1998; 98US-0101738P.
XX 01-SEP-1998; 98US-0101741P.
XX 01-SEP-1998; 98US-0101743P.
XX 01-SEP-1998; 98US-0101915P.
XX 01-SEP-1998; 98US-0101916P.
XX 01-SEP-1998; 98US-0102207P.
XX 01-SEP-1998; 98US-0102240P.
XX 01-SEP-1998; 98US-0102307P.
XX 01-SEP-1998; 98US-0102331P.
XX 01-SEP-1998; 98US-0102484P.
XX 01-SEP-1998; 98US-0102487P.
XX 01-SEP-1998; 98US-0102570P.
XX 01-SEP-1998; 98US-0102571P.
XX 01-SEP-1998; 98US-0102584P.
XX 01-SEP-1998; 98US-0102687P.
XX 01-SEP-1998; 98US-0102965P.
XX 01-SEP-1998; 98US-0103258P.
XX 01-SEP-1998; 98US-0103449P.
XX 01-SEP-1998; 98US-0103314P.
XX 01-SEP-1998; 98US-0103315P.
XX 01-SEP-1998; 98US-0103328P.
XX 01-SEP-1998; 98US-0103395P.
XX 01-SEP-1998; 98US-0103396P.
XX 01-SEP-1998; 98US-0103401P.
XX 01-SEP-1998; 98US-0103633P.
XX 01-SEP-1998; 98US-0103678P.
XX 01-SEP-1998; 98US-0103679P.
XX 01-SEP-1998; 98US-0103711P.
XX 01-SEP-1998; 98US-0104257P.
XX 01-SEP-1998; 98US-0104987P.
XX 01-SEP-1998; 98US-0105000P.
XX 01-SEP-1998; 98US-0105002P.
XX 01-SEP-1998; 98US-0105104P.
XX 01-SEP-1998; 98US-0105169P.
XX 01-SEP-1998; 98US-0105266P.
XX 01-SEP-1998; 98US-0105893P.
XX 01-SEP-1998; 98US-0105894P.
XX 01-SEP-1998; 98US-0105807P.
XX 01-SEP-1998; 98US-0105881P.
XX 01-SEP-1998; 98US-0105882P.
XX 01-SEP-1998; 98US-0106062P.
XX 01-SEP-1998; 98US-0106023P.
XX 01-SEP-1998; 98US-0106029P.
XX 01-SEP-1998; 98US-0106030P.
XX 01-SEP-1998; 98US-0106032P.
XX 01-SEP-1998; 98US-0106033P.
XX 01-SEP-1998; 98US-0106178P.
XX 01-SEP-1998; 98US-0106248P.
XX 01-SEP-1998; 98US-0106384P.
XX 01-SEP-1998; 98US-0106500P.
XX 01-SEP-1998; 98US-0106464P.
XX 01-SEP-1998; 98US-0106856P.
XX 01-SEP-1998; 98US-0106902P.
```


Db 961 ACTCTTTGGCTGGGAAGCTTCTTGAACCTTTAACTCTCGCAGCCCTTCTAAGACCCA 1020
QY 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATATATAAATGAAGGGGCAAA 1080
Db 1021 CGAGCGGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATATATAAATGAAGGGGCAAA 1080
QY 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091

RESULT 10

ADD39785
ID ADD39785 standard; cDNA; 1091 BP.

XX
AC ADD39785;

DT 15-JAN-2004 (first entry)

XX Human cDNA encoding secreted/transmembrane protein PRO1303.

XX Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW dermatitis; herpetiformis; Crohn's disease; thalassaemia.

XX Homo sapiens.

OS
PN US2003083462-Al.

XX
XX 01-MAY-2003.

PD
PF 10-DEC-2001; 2001US-00013913.

XX
XX 05-JAN-1999; 99WO-US000106.

PR 01-SEP-1999; 99WO-US020111.

PR 15-SEP-1999; 99WO-US021134.

PR 30-NOV-1999; 99WO-US028313.

PR 02-DEC-1999; 99WO-US028551.

PR 16-DEC-1999; 99WO-US030095.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000376.

PR 11-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004342.

PR 24-FEB-2000; 2000WO-US005004.

PR 02-MAR-2000; 2000WO-US005841.

PR 15-MAR-2000; 2000WO-US006884.

PR 17-MAY-2000; 2000WO-US013705.

PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.

PR 23-AUG-2000; 2000WO-US023522.

PR 24-AUG-2000; 2000WO-US023328.

PR 08-NOV-2000; 2000WO-US030952.

PR 10-NOV-2000; 2000WO-US030873.

PR 01-DEC-2000; 2000WO-US032678.

PR 28-FEB-2001; 2001WO-US006520.

PR 01-MAR-2001; 2001WO-US006666.

PR 01-JUN-2001; 2001WO-US017800.

PR 20-JUN-2001; 2001WO-US019692.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-JUL-2001; 2001WO-US021735.

PR 04-SEP-2001; 2001US-00946374.

XX
PA (GETH) GENENTECH INC.

XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;

XX

DR WPI; 2003-755122/71.
DR P-PSDB; ADD39786.

XX

DR

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 1091 BP; 221 A; 335 C; 313 G; 222 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 1091; DB 9; Length 1091;

Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

```
Db 361 GACTGGACCGAGCAGATCCGACAGCGGCTTCTGTGACCCATCCCGGCTACCTGGGA 420
Qy 421 GCCTCGAGCGACGACGAGCAGCTCCGGTGTCTGGGCTGGGCTGCCGTCGGCGTA 480
Db 421 GCTTCGAGCGACGACGAGCAGCTCCGGTGTCTGGGCTGGGCTGCCGTCGGCGTA 480
Qy 481 ACCAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGC 540
Db 481 ACCAGCAGCGTTCAACCCCTGCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGC 540
Qy 541 CACGTCTCAGCTGGGGATCACCAACACCGAACCATTCCCGATCTCTCCAG 600
Db 541 CACGTCTCAGCTGGGGATCACCAACACCGAACCATTCCCGATCTCTCCAG 600
Qy 601 TGCCTCAACCTCTCCATGCTCCCATGCCACCTGCGCATGGTGTATCCCGGAGATC 660
Db 601 TGCCTCAACCTCTCCATGCTCCCATGCCACCTGCGCATGGTGTATCCCGGAGATC 660
Qy 661 ACAGCAACATGTGTGTGACAGCGGGGTCCCGGGGAGGATGCTGCCAGGGTGATTCT 720
Db 661 ACAGCAACATGTGTGTGACAGCGGGGTCCCGGGGAGGATGCTGCCAGGGTGATTCT 720
Qy 721 GGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGGTCTGTGTCTCGGGGTCTGTGGGG 780
Db 721 GGGGGCCCCCTGTGTGTGGGGAGTCTTCAAGGTCTGTGTCTCGGGGTCTGTGGGG 780
Qy 781 CCTGTGGACAAGATGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840
Db 781 CCTGTGGACAAGATGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGGACTGG 840
Qy 841 ATCCGATGATCATGAGGAACAATGACCTGTCTTCCCTCCACCTCCACCCCAACCCCTTAA 900
Db 841 ATCCGATGATCATGAGGAACAATGACCTGTCTTCCCTCCACCTCCACCCCAACCCCTTAA 900
Qy 901 CTTGGGTACCCCTTGGCCCTGAGACCAATATCTCTCCATCTCTCCCTAGCTCC 960
Db 901 CTTGGGTACCCCTTGGCCCTGAGACCAATATCTCTCCATCTCTCCCTAGCTCC 960
Qy 961 ACTCTTGTGGCTGGGAATCTTGGAACTTTTAACTCTCCGAGCCCTCTAAGACCCA 1020
Db 961 ACTCTTGTGGCTGGGAATCTTGGAACTTTTAACTCTCCGAGCCCTCTAAGACCCA 1020
Qy 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAAGGAGGGGCAAA 1080
Db 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAAGGAGGGGCAAA 1080
Qy 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091

RESULT 11
ADD70231
ID ADD70231 standard; cdna; 1091 BP.
XX
AC ADD70231;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human cdna encoding secreted/transmembrane protein PRO1303.
XX
KW Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX
OS Homo sapiens.
XX
FN US2003054406-A1.
XX
PD 20-MAR-2003.

06-DEC-2001; 2001US-00006818.
01-SEP-1998; 98US-0098716P.
01-SEP-1998; 98US-0098723P.
01-SEP-1998; 98US-0098749P.
01-SEP-1998; 98US-0098750P.
02-SEP-1998; 98US-0098803P.
02-SEP-1998; 98US-0098821P.
02-SEP-1998; 98US-0098843P.
02-SEP-1998; 98US-0098953P.
09-SEP-1998; 98US-0099536P.
09-SEP-1998; 98US-0099598P.
09-SEP-1998; 98US-0099602P.
09-SEP-1998; 98US-0099642P.
09-SEP-1998; 98US-0099741P.
10-SEP-1998; 98US-0099754P.
10-SEP-1998; 98US-0099763P.
10-SEP-1998; 98US-0099792P.
10-SEP-1998; 98US-0099808P.
10-SEP-1998; 98US-0099812P.
10-SEP-1998; 98US-0099815P.
10-SEP-1998; 98US-0099816P.
15-SEP-1998; 98US-0100385P.
15-SEP-1998; 98US-0100388P.
15-SEP-1998; 98US-0100390P.
16-SEP-1998; 98US-0100584P.
16-SEP-1998; 98US-0100627P.
16-SEP-1998; 98US-0100651P.
16-SEP-1998; 98US-0100661P.
16-SEP-1998; 98US-0100662P.
17-SEP-1998; 98US-0100664P.
17-SEP-1998; 98US-0100683P.
17-SEP-1998; 98US-0100684P.
17-SEP-1998; 98US-0100710P.
17-SEP-1998; 98US-0100711P.
17-SEP-1998; 98US-0100919P.
17-SEP-1998; 98US-0100930P.
18-SEP-1998; 98US-0100948P.
18-SEP-1998; 98US-0100949P.
18-SEP-1998; 98US-0101014P.
18-SEP-1998; 98US-0101068P.
18-SEP-1998; 98US-0101071P.
22-SEP-1998; 98US-0101279P.
23-SEP-1998; 98US-0101471P.
23-SEP-1998; 98US-0101472P.
23-SEP-1998; 98US-0101474P.
23-SEP-1998; 98US-0101475P.
23-SEP-1998; 98US-0101476P.
23-SEP-1998; 98US-0101477P.
23-SEP-1998; 98US-0101479P.
24-SEP-1998; 98US-0101738P.
24-SEP-1998; 98US-0101741P.
24-SEP-1998; 98US-0101743P.
24-SEP-1998; 98US-0101915P.
24-SEP-1998; 98US-0101916P.
29-SEP-1998; 98US-010207P.
29-SEP-1998; 98US-0102240P.
29-SEP-1998; 98US-0102307P.
29-SEP-1998; 98US-0102331P.
30-SEP-1998; 98US-0102484P.
30-SEP-1998; 98US-0102487P.
30-SEP-1998; 98US-0102570P.
01-OCT-1998; 98US-0102571P.
01-OCT-1998; 98US-0102684P.
02-OCT-1998; 98US-0102687P.
06-OCT-1998; 98US-0102965P.
06-OCT-1998; 98US-0103258P.
07-OCT-1998; 98US-0103449P.
07-OCT-1998; 98US-0103314P.
07-OCT-1998; 98US-0103315P.
07-OCT-1998; 98US-0103328P.
07-OCT-1998; 98US-0103395P.
```



```
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105593P.
PR 26-OCT-1998; 98US-0105594P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 08-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.

PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 04-SEP-2001; 2001US-00946374.
XX (GETH ) GENENTECH INC.
PA Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,
XX Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;
XX WPI; 2003-787000/74.
DR P-PSDB; AD038353.
XX
XX Novel isolated PRO polypeptide, useful for treating cancerous tumors,
PT cardiac insufficiency disorders, wound healing, diabetes mellitus,
PT thalassemias.
XX
XX Claim 2; SEQ ID NO 193; 556pp; English.
PS
CC The invention relates to an isolated PRO polypeptide (secreted or
CC transmembrane protein) having at least 80% amino acid sequence identity
Query Match 100.0%; Score 1091; DB 9; Length 1091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCAGGTCATCCCTCGTGCACCTTCAAGAGAGACAGAGAGGACAGAGTGGGGGC 60
DB 1 CAAGCAGGTCATCCCTCGTGCACCTTCAAGAGAGAGACAGAGAGGACAGAGTGGGGGC 60
QY 61 ACAGGGAAGAGGAGTCTCTGAGATTCCCTTTTCCCTCCAGACTTTTGGAGTGACCCACC 120
DB 61 ACAGGGAAGAGGAGTCTCTGAGATTCCCTTTTCCCTCCAGACTTTTGGAGTGACCCACC 120
QY 121 ATGGGGCTCAGCATCTTTTGTCTCTCTGTGTCTTCTGGGCTCAGCAGGACGACCG 180
DB 121 ATGGGGCTCAGCATCTTTTGTCTCTCTGTGTCTTCTGGGCTCAGCAGGACGACCG 180
QY 181 AAGATTTTCAATGGCACTGAGTGTGGCGTAACTCACTCAGCGTGGCAGGTGGGGCTGTTT 240
DB 181 AAGATTTTCAATGGCACTGAGTGTGGCGTAACTCACTCAGCGTGGCAGGTGGGGCTGTTT 240
QY 241 GAGGGCACCAAGCTCGCGTGGGGGTGTCCTTATTGACACAGTGGGTCTTCAAGCG 300
DB 241 GAGGGCACCAAGCTCGCGTGGGGGTGTCCTTATTGACACAGTGGGTCTTCAAGCG 300
QY 301 GCTCACTGAGCGGACAGGTAAGTCTGGGTGGCGCTGGGGGAACACAGCTCAGCCAGTTC 360
DB 301 GCTCACTGAGCGGACAGGTAAGTCTGGGTGGCGCTGGGGGAACACAGCTCAGCCAGTTC 360
QY 361 GACTGGACCGAGCAGATCCGGGCACAGCGGCTTCTGTGTGACCATCCCGCTACTTGGGA 420
DB 361 GACTGGACCGAGCAGATCCGGGCACAGCGGCTTCTGTGTGACCATCCCGCTACTTGGGA 420
QY 421 GCCTCGAGGACCAAGCAGCAGCTCCGGCTGTCTGGGTGGCGCTGGCGCTGCCCGTCCCGTA 480
DB 421 GCCTCGAGGACCAAGCAGCAGCTCCGGCTGTCTGGGTGGCGCTGGCGCTGCCCGTCCCGTA 480
QY 481 ACCAGCAGCGTTCAACCCCTGCCCTGCGCAATGACTGTGCAACCGCTGSCACCGAGTGC 540
DB 481 ACCAGCAGCGTTCAACCCCTGCCCTGCGCAATGACTGTGCAACCGCTGSCACCGAGTGC 540
QY 541 CAGCTCTCAGGCTGGGGCATCAACAAACCCACCGAAACCCATTCGGGATCTGTCCAG 600
DB 541 CAGCTCTCAGGCTGGGGCATCAACAAACCCACCGAAACCCATTCGGGATCTGTCCAG 600
QY 601 TGCTCAACCTCTCCATCGTCTCCATGCGCACTGCGCATGTGTGTATCTCCGGGAGATC 660
DB 601 TGCTCAACCTCTCCATCGTCTCCATGCGCACTGCGCATGTGTGTATCTCCGGGAGATC 660
```

```
Db      601  TGCCTCAACCTCTCATCGTCTCCATGCCACCTGCCATGGTGTATATCCCGGAGATC 660
Qy      661  ACGAGCAACATGGTGTGTCAGGCGGCTCCCGGCGAGGATGCTGCCAGGGTATTCT 720
Db      661  ACGAGCAACATGGTGTGTCAGGCGGCTCCCGGCGAGGATGCTGCCAGGGTATTCT 720
Qy      721  GGGGGCCCCCTGGTGTGTTGGGGGAGTCCTTCAAGTCTGGTGTCTGGGGGTCTGTGGG 780
Db      721  GGGGGCCCCCTGGTGTGTTGGGGGAGTCCTTCAAGTCTGGTGTCTGGGGGTCTGTGGG 780
Qy      781  CCCTGTGACAAAGATGGCATCCCTGGAGTCTACACTATATTTGCAAGTATGTGACTGG 840
Db      781  CCCTGTGACAAAGATGGCATCCCTGGAGTCTACACTATATTTGCAAGTATGTGACTGG 840
Qy      841  ATCCGGATGATCATGAGAAACAACATGACCTGTTCTCTCCACCTCCACCCCAACCCCTTAA 900
Db      841  ATCCGGATGATCATGAGAAACAACATGACCTGTTCTCTCCACCTCCACCCCAACCCCTTAA 900
Qy      901  CTTGGGTACCCCTCTGGCCCTCAGAGACCAATATCTCTCCATCACTTCCCTAGCTCC 960
Db      901  CTTGGGTACCCCTCTGGCCCTCAGAGACCAATATCTCTCCATCACTTCCCTAGCTCC 960
Qy      961  ACTCTTTGGCTCGGAACTTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Db      961  ACTCTTTGGCTCGGAACTTCTTGGAACTTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Qy      1021  CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAAATAAATAAATGAAGGGGGCAAA 1080
Db      1021  CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAAATAAATAAATGAAGGGGGCAAA 1080
Qy      1081  AAAAAAAAAA 1091
Db      1081  AAAAAAAAAA 1091

RESULT 13
ADD33308
ID  ADD33308 standard; cDNA; 1091 BP.
AC  ADD33308;
XX
XX
DT  15-JAN-2004 (first entry)
XX
DE  Human cDNA encoding secreted/transmembrane protein PRO1303.
XX
KW  Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW  immune response; cardiac insufficiency disorder; calcium flux;
KW  umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW  arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW  Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW  dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX
OS  Homo sapiens.
XX
XX  US2003096954-A1.
PN  XX
PD  22-MAY-2003.
XX
XX  07-DEC-2001; 2001US-00011671.
PF  XX
XX  01-SEP-1998; 98US-0098716P.
PR  01-SEP-1998; 98US-0098723P.
PR  01-SEP-1998; 98US-0098749P.
PR  01-SEP-1998; 98US-0098750P.
PR  02-SEP-1998; 98US-0098803P.
PR  02-SEP-1998; 98US-0098821P.
PR  02-SEP-1998; 98US-0098843P.
PR  09-SEP-1998; 98US-0099536P.
PR  09-SEP-1998; 98US-0099596P.
PR  09-SEP-1998; 98US-0099602P.
PR  09-SEP-1998; 98US-0099642P.
PR  10-SEP-1998; 98US-0099741P.
PR  10-SEP-1998; 98US-0099754P.
PR  10-SEP-1998; 98US-0099763P.
PR  10-SEP-1998; 98US-0099792P.
PR  10-SEP-1998; 98US-0099808P.
PR  10-SEP-1998; 98US-0099812P.
PR  10-SEP-1998; 98US-0099815P.
PR  10-SEP-1998; 98US-0099816P.
PR  15-SEP-1998; 98US-0100385P.
PR  15-SEP-1998; 98US-0100388P.
PR  15-SEP-1998; 98US-0100390P.
PR  16-SEP-1998; 98US-0100584P.
PR  16-SEP-1998; 98US-0100627P.
PR  16-SEP-1998; 98US-0100661P.
PR  16-SEP-1998; 98US-0100662P.
PR  16-SEP-1998; 98US-0100664P.
PR  17-SEP-1998; 98US-0100683P.
PR  17-SEP-1998; 98US-0100684P.
PR  17-SEP-1998; 98US-0100710P.
PR  17-SEP-1998; 98US-0100711P.
PR  17-SEP-1998; 98US-0100919P.
PR  17-SEP-1998; 98US-0100930P.
PR  18-SEP-1998; 98US-0100848P.
PR  18-SEP-1998; 98US-0100849P.
PR  18-SEP-1998; 98US-0101014P.
PR  18-SEP-1998; 98US-0101068P.
PR  18-SEP-1998; 98US-0101071P.
PR  22-SEP-1998; 98US-0101279P.
PR  23-SEP-1998; 98US-0101471P.
PR  23-SEP-1998; 98US-0101472P.
PR  23-SEP-1998; 98US-0101474P.
PR  23-SEP-1998; 98US-0101475P.
PR  23-SEP-1998; 98US-0101476P.
PR  23-SEP-1998; 98US-0101477P.
PR  23-SEP-1998; 98US-0101479P.
PR  24-SEP-1998; 98US-0101738P.
PR  24-SEP-1998; 98US-0101741P.
PR  24-SEP-1998; 98US-0101743P.
PR  24-SEP-1998; 98US-0101916P.
PR  29-SEP-1998; 98US-0102207P.
PR  29-SEP-1998; 98US-0102240P.
PR  29-SEP-1998; 98US-0102307P.
PR  29-SEP-1998; 98US-0102330P.
PR  29-SEP-1998; 98US-0102331P.
PR  30-SEP-1998; 98US-0102484P.
PR  30-SEP-1998; 98US-0102487P.
PR  30-SEP-1998; 98US-0102570P.
PR  30-SEP-1998; 98US-0102571P.
PR  01-OCT-1998; 98US-0102684P.
PR  01-OCT-1998; 98US-0102687P.
PR  02-OCT-1998; 98US-0102965P.
PR  06-OCT-1998; 98US-0103258P.
PR  06-OCT-1998; 98US-0103314P.
PR  07-OCT-1998; 98US-0103315P.
PR  07-OCT-1998; 98US-0103328P.
PR  07-OCT-1998; 98US-0103355P.
PR  07-OCT-1998; 98US-0103396P.
PR  08-OCT-1998; 98US-0103401P.
PR  08-OCT-1998; 98US-0103633P.
PR  08-OCT-1998; 98US-0103678P.
PR  08-OCT-1998; 98US-0103679P.
PR  14-OCT-1998; 98US-0103711P.
PR  20-OCT-1998; 98US-0104987P.
PR  20-OCT-1998; 98US-0105000P.
PR  20-OCT-1998; 98US-0105002P.
PR  21-OCT-1998; 98US-0105104P.
PR  22-OCT-1998; 98US-0105169P.
PR  22-OCT-1998; 98US-0105266P.
PR  26-OCT-1998; 98US-0105693P.
PR  26-OCT-1998; 98US-0105694P.
PR  27-OCT-1998; 98US-0105807P.
```



```
Db      721 GGGGGCCCCCTGGTGTGGGGGAGTCCTTCAAGGTGTGGTCTCTGGGGTCTGTGGG 780
Qy      781 CCCTGTGGAAGAATGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGACATGG 840
Db      781 CCCTGTGGAAGAATGGCATCCCTGGAGTCTACACCTATATTTGCAAGTATGTGACATGG 840
Qy      841 ATCCGGATGATCATGAGAAACAACCTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900
Db      841 ATCCGGATGATCATGAGAAACAACCTGACCTGTTTCTCCACCTCCACCCCAACCCCTTAA 900
Qy      901 CTTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCTCCATCACATCCCTAGTCCTC 960
Db      901 CTTGGGTACCCCTCTGGGCCCTCAGAGCAACCAATATCTCTCCATCACATCCCTAGTCCTC 960
Qy      961 ACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTCTGCGAGCCCTTCTAGACCCA 1020
Db      961 ACTCTTGTGGCTGGGAACCTTCTTGGAACTTTAACTCTCTGCGAGCCCTTCTAGACCCA 1020
Qy      1021 CGAGCGGGTGCAGAACTGTGCAATAGTCTGGAATAATATAATGAAGGAGGCAAA 1080
Db      1021 CGAGCGGGTGCAGAACTGTGCAATAGTCTGGAATAATATAATGAAGGAGGCAAA 1080
Qy      1081 AAAAAAAAAA 1091
Db      1081 AAAAAAAAAA 1091

RESULT 14
ADD38831
ID      ADD38831 standard; cDNA; 1091 BP.
XX
AC      ADD38831;
XX
DT      15-JAN-2004 (first entry)
XX
DE      Human cDNA encoding secreted/transmembrane protein PRO1303.
KW      Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW      immune response; cardiac insufficiency disorder; calcium flux;
KW      umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW      arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW      Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW      dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX
OS      Homo sapiens.
XX
FN      US2003092061-A1.
XX
PD      15-MAY-2003.
XX
PF      06-DEC-2001; 2001US-00007194.
XX
PR      01-SEP-1998; 98US-0038716P.
PR      01-SEP-1998; 98US-0038723P.
PR      01-SEP-1998; 98US-0038749P.
PR      01-SEP-1998; 98US-0038750P.
PR      02-SEP-1998; 98US-0038803P.
PR      02-SEP-1998; 98US-0038821P.
PR      02-SEP-1998; 98US-0038843P.
PR      09-SEP-1998; 98US-0039536P.
PR      09-SEP-1998; 98US-0039596P.
PR      09-SEP-1998; 98US-0039602P.
PR      09-SEP-1998; 98US-0039642P.
PR      10-SEP-1998; 98US-0039741P.
PR      10-SEP-1998; 98US-0039763P.
PR      10-SEP-1998; 98US-0039792P.
PR      10-SEP-1998; 98US-0039808P.
PR      10-SEP-1998; 98US-0039812P.
PR      10-SEP-1998; 98US-0039815P.
PR      10-SEP-1998; 98US-0039816P.
PR      15-SEP-1998; 98US-0100385P.
PR      15-SEP-1998; 98US-0100388P.
PR      15-SEP-1998; 98US-0100390P.
PR      16-SEP-1998; 98US-0100584P.
PR      16-SEP-1998; 98US-0100627P.
PR      16-SEP-1998; 98US-0100661P.
PR      16-SEP-1998; 98US-0100682P.
PR      16-SEP-1998; 98US-0100684P.
PR      17-SEP-1998; 98US-0100683P.
PR      17-SEP-1998; 98US-0100710P.
PR      17-SEP-1998; 98US-0100711P.
PR      17-SEP-1998; 98US-0100919P.
PR      17-SEP-1998; 98US-0100930P.
PR      18-SEP-1998; 98US-0100848P.
PR      18-SEP-1998; 98US-0100849P.
PR      18-SEP-1998; 98US-0101014P.
PR      18-SEP-1998; 98US-0101068P.
PR      18-SEP-1998; 98US-0101071P.
PR      22-SEP-1998; 98US-0101279P.
PR      23-SEP-1998; 98US-0101471P.
PR      23-SEP-1998; 98US-0101472P.
PR      23-SEP-1998; 98US-0101474P.
PR      23-SEP-1998; 98US-0101475P.
PR      23-SEP-1998; 98US-0101476P.
PR      23-SEP-1998; 98US-0101477P.
PR      23-SEP-1998; 98US-0101479P.
PR      24-SEP-1998; 98US-0101738P.
PR      24-SEP-1998; 98US-0101741P.
PR      24-SEP-1998; 98US-0101743P.
PR      24-SEP-1998; 98US-0101915P.
PR      24-SEP-1998; 98US-0101916P.
PR      29-SEP-1998; 98US-0102207P.
PR      29-SEP-1998; 98US-0102240P.
PR      29-SEP-1998; 98US-0102307P.
PR      29-SEP-1998; 98US-0102330P.
PR      30-SEP-1998; 98US-0102331P.
PR      30-SEP-1998; 98US-0102484P.
PR      30-SEP-1998; 98US-0102487P.
PR      30-SEP-1998; 98US-0102570P.
PR      01-OCT-1998; 98US-0102571P.
PR      01-OCT-1998; 98US-0102684P.
PR      02-OCT-1998; 98US-0102687P.
PR      06-OCT-1998; 98US-0102965P.
PR      06-OCT-1998; 98US-0103258P.
PR      07-OCT-1998; 98US-0103449P.
PR      07-OCT-1998; 98US-0103314P.
PR      07-OCT-1998; 98US-0103315P.
PR      07-OCT-1998; 98US-0103328P.
PR      07-OCT-1998; 98US-0103395P.
PR      07-OCT-1998; 98US-0103396P.
PR      07-OCT-1998; 98US-0103401P.
PR      08-OCT-1998; 98US-0103633P.
PR      08-OCT-1998; 98US-0103678P.
PR      08-OCT-1998; 98US-0103679P.
PR      14-OCT-1998; 98US-0103711P.
PR      20-OCT-1998; 98US-0104257P.
PR      20-OCT-1998; 98US-0104987P.
PR      20-OCT-1998; 98US-0105000P.
PR      21-OCT-1998; 98US-0105104P.
PR      22-OCT-1998; 98US-0105169P.
PR      22-OCT-1998; 98US-0105266P.
PR      26-OCT-1998; 98US-0105693P.
PR      26-OCT-1998; 98US-0105694P.
PR      27-OCT-1998; 98US-0105807P.
PR      27-OCT-1998; 98US-0105881P.
PR      27-OCT-1998; 98US-0105882P.
PR      28-OCT-1998; 98US-0106023P.
PR      28-OCT-1998; 98US-0106029P.
PR      28-OCT-1998; 98US-0106030P.
PR      28-OCT-1998; 98US-0106032P.
PR      28-OCT-1998; 98US-0106033P.
```

```
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0106500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 05-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 04-SEP-2001; 2001US-00946374.
PR XX
PR (GETH ) GENENTECH INC.
PR
PR Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PR Gao W, Goddard A, Godowski EJ, Grimaldi JC, Gurney AL, Hillan KJ;
PR Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CX;
PR Williams PM, Wood WI;
PR XX
XX
```

Db 841 ATCCGGATGATCATGAGGACAACTGACCTGTTTCTCCACCTCCACCCACCCCTTAA 900
QY 901 CTGTGGTACCCCTCTGGCCCTCAGAGACCAATATCTCTCCATCACTTCCCTAGCTCC 960
Db 901 CTGTGGTACCCCTCTGGCCCTCAGAGACCAATATCTCTCCATCACTTCCCTAGCTCC 960
QY 961 ACTCTGTGGCTGGGAACTTCTTGGAACTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
Db 961 ACTCTGTGGCTGGGAACTTCTTGGAACTTAACTCTGCGAGCCCTTCTAAGACCCA 1020
QY 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAATAAATAAATAA 1080
Db 1021 CGAGCGGGTGAGAGAGTGTGCAATAGTCTGGAATAAATAAATAAATAAATAAATAA 1080
QY 1081 AAAAAAAAAA 1091
Db 1081 AAAAAAAAAA 1091
RESULT 15
ADD40262
ID ADD40262 standard; cDNA; 1091 BP.
XX AC
AC ADD40262;
XX AC
DT 15-JAN-2004 (first entry)
DE Human cDNA encoding secreted/transmembrane protein PRO1303.
XX Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonlein-Henoch purpura; colliac disease;
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX Homo sapiens.
OS
XX
XX
XX US2003082627-A1.
PD
XX
XX 01-MAY-2003.
XX
XX 06-DEC-2001; 2001US-00006117.
PF
XX
XX 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 01-SEP-1998; 98US-0098749P.
PR 01-SEP-1998; 98US-0098750P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099536P.
PR 09-SEP-1998; 98US-0099596P.
PR 09-SEP-1998; 98US-0099598P.
PR 09-SEP-1998; 98US-0099602P.
PR 09-SEP-1998; 98US-0099642P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099792P.
PR 10-SEP-1998; 98US-0099808P.
PR 10-SEP-1998; 98US-0099812P.
PR 10-SEP-1998; 98US-0099815P.
PR 10-SEP-1998; 98US-0099816P.
PR 15-SEP-1998; 98US-0100385P.
PR 15-SEP-1998; 98US-0100388P.
PR 15-SEP-1998; 98US-0100390P.
PR 16-SEP-1998; 98US-0100584P.
PR 16-SEP-1998; 98US-0100627P.
PR 16-SEP-1998; 98US-0100661P.
PR 16-SEP-1998; 98US-0100662P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100710P.
PR 17-SEP-1998; 98US-0100711P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100848P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 18-SEP-1998; 98US-0101071P.
PR 22-SEP-1998; 98US-0101279P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101474P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101476P.
PR 23-SEP-1998; 98US-0101477P.
PR 23-SEP-1998; 98US-0101479P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 20-OCT-1998; 98US-0103711P.
PR 20-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 21-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 29-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0108464P.
PR 03-NOV-1998; 98US-0108565P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.

[illegible]

Db	961	ACTCTTTGTCGCTGGGNACTTCTTGGA	CTTTAACTCCTGCCAGCCCTTCTAAGACCCA	1020
Qy	1021	CGAGCGGGTGAGAGAACTGTGCAATAGTCTGGA	ATAAATAATGAAGAGGGGCAAA	1080
Db	1021	CGAGCGGGTGAGAGAACTGTGCAATAGTCTGGA	ATAAATAATGAAGAGGGGCAAA	1080
Qy	1081	AAAAAAAAAA	1091	
Db	1081	AAAAAAAAAA	1091	

Search completed: June 20, 2004, 06:17:42
Job time : 512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2004, 06:09:07 ; Search time 3353 Seconds
(without alignments)
9716.564 Million cell updates/sec

Title: US-10-015-385A-193
Perfect score: 1091
Sequence: 1 caagcaggtcatcccttgg.....aggggcaaaaaaaaaa 1091

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmd.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gsei.*
29: gb_gses.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	325	29.8	472	13	BX110362
C 2	254	23.3	469	9	AL394679
3	250	22.9	1186	11	BC035385
4	216	19.8	264	12	BG230967

5	101	9.3	383	9	AI554072
C 6	101	9.3	478	13	BX103516
7	52	4.8	550	12	BG721150
C 8	26	2.4	577	13	BQ129094
9	25	2.3	600	13	BQ270082
10	25	2.3	636	13	BQ269009
C 11	24	2.2	283	12	BQ794700
12	23	2.1	449	9	AI322408
C 13	23	2.1	451	14	CA313295
14	23	2.1	506	9	AA028356
15	23	2.1	520	13	BX514743
C 16	23	2.1	545	9	AI324852
17	23	2.1	627	29	CE842883
18	23	2.1	763	14	CD646427
19	23	2.1	765	12	BG762809
20	23	2.1	768	29	CC558299
21	23	2.1	890	11	AK009217
22	23	2.1	898	13	BY709025
C 23	23	2.1	904	13	EX689949
C 24	23	2.1	1002	29	CNS05FCB
C 25	22	2.0	149	12	BI278918
C 26	22	2.0	224	9	AI924035
C 27	22	2.0	242	10	BF353225
C 28	22	2.0	255	29	CE737502
29	22	2.0	287	12	BM368334
C 30	22	2.0	351	10	BF073014
31	22	2.0	389	12	BI346787
32	22	2.0	394	14	CB776687
33	22	2.0	400	9	AA293231
C 34	22	2.0	402	10	AM136489
35	22	2.0	405	12	BM760479
C 36	22	2.0	457	10	BE711073
C 37	22	2.0	464	14	CF308398
C 38	22	2.0	482	10	AW820357
C 39	22	2.0	483	10	AW994056
C 40	22	2.0	491	14	CB725454
41	22	2.0	498	14	W73168
42	22	2.0	512	12	BM830283
C 43	22	2.0	515	14	CB716451
44	22	2.0	517	9	AA293027
45	22	2.0	523	12	BM841293

ALIGNMENTS

RESULT 1
BX110362
LOCUS
DEFINITION BX110362 NCI CGAP CLL1 Homo sapiens CDNA clone IMAGE:2109677, mRNA sequence.
IMAGE:2109677, mRNA sequence.
BX110362
ACCESSION BX110362.1 GI:27836332
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGE998E065203
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAACAGCTATGAC.

FEATURES

source

1. .472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:988065203 ; IMAGE:2109677"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI CGAP CLL1"
/notes="Vector: p7T73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5],
TGTACCAATCTGAAGTGGAGCGCGCATCTGCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 29.8%; Score 325; DB 13; Length 472;
Best Local Similarity 99.7%; Pred. No. 2e-82;
Matches 375; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 100 AGACTTTGGAAGTCAACCCAGCCAGTGGGCTCAGCATCTTTTGTCCCTGCTGTTCTTGGG 159
DB 97 AGACTTTGGAAGTCAACCCAGCCAGTGGGCTCAGCATCTTTTGTCCCTGCTGTTCTTGGG 156
QY 160 CTCAGCAGGAGCCACACAGCAAGATTTCATGGCACTGAGTGTGGCGCTAACTACAG 219
DB 157 CTCAGCAGGAGCCACACAGCAAGATTTCATGGCACTGAGTGTGGCGCTAACTACAG 216
QY 220 CCGTGGCAGTGGGGCTGTTGAGGGACACAGCTGCTGCGGGGGTGTCTTATTGAC 279
DB 217 CCGTGGCAGTGGGGCTGTTGAGGGACACAGCTGCTGCGGGGGTGTCTTATTGAC 276
QY 280 CACAGTGGGTCTCTACAGCGGCTCACTGACGCGGAGCAGAGTACTGGGTGCGCTCGGG 339
DB 277 CACAGTGGGTCTCTACAGCGGCTCACTGACGCGGAGCAGAGTACTGGGTGCGCTCGGG 336
QY 340 GAACACAGCTCAACCGAGCTCACTGACCGGAGCAGATCCGGCAACAGCGGCTTCTCTGTG 399
DB 337 GAACACAGCTCAACCGAGCTCACTGACCGGAGCAGATCCGGCAACAGCGGCTTCTCTGTG 396
QY 400 ACCATCCCGCTACCTGGGAGCCTCGACGAGCCACAGCAGCAGCTCCGCTCGCTCGGG 459
DB 397 ACCATCCCGCTACCTGGGAGCCTCGACGAGCCACAGCAGCAGCTCCGCTCGCTCGGG 456
QY 460 CTGGCGCTGCGCGTCC 475
DB 457 CTGGCGCTGCGCGTCC 472

RESULT 2

AI394679/c

LOCUS

AI394679 469 bp mRNA linear EST 30-MAR-1999
tq24b03.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2109677 3'
Similar to SW:TRY3_CHICK Q90629 TRYPSINOGEN 2-P29 PRECURSOR.
; contains TAR1.D1 TAR1 repetitive element ; mRNA sequence.

ACCESSION

AI394679

VERSION

AI394679.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

1 (bases 1 to 469)

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,

M.D., Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

cDNA Library Arrayed by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1051 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 397.

Location/Qualifiers

source

1. .469

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2109677"

/tissue_type="B-cell, chronic lymphocytic leukemia"

/lab_host="DH10B"

/clone_lib="NCI CGAP CLL1"

/notes="Vector: p7T73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5],

TGTACCAATCTGAAGTGGAGCGCGCATCTGCTTTTCTTTTCTTTTCTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified p7T73 vector.

Library is normalized, and was constructed by Bento

Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 23.3%; Score 254; DB 9; Length 469;

Best Local Similarity 100.0%; Pred. No. 2.5e-62;

Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 CGGAACCCATTCCTCCGATCTGCTCCAGTGCCTCAACCTCTCCATCGTCTCCATGCCACC 633
DB 390 CGGAACCCATTCCTCCGATCTGCTCCAGTGCCTCAACCTCTCCATCGTCTCCATGCCACC 331
QY 634 TGCATGTGTGTATCCCGGAGAAATCAGAGCAACATGTGTGTGTCAGCGCGGCTCCCG 693
DB 330 TGCATGTGTGTATCCCGGAGAAATCAGAGCAACATGTGTGTGTCAGCGCGGCTCCCG 271
QY 694 GGCAGGATGCCCTGCCAGGGTGTCTGGGGGCCCCCTGTGTGTGGGGAGTCTTCAA 753
DB 270 GGCAGGATGCCCTGCCAGGGTGTCTGGGGGCCCCCTGTGTGTGGGGAGTCTTCAA 211
QY 754 GGTCTGTGTCTGGGGGTCTGTGGGGCCCCCTGTGGGCAAGATGGCATCCCTGGAGTCTAC 813
DB 210 GGTCTGTGTCTGGGGGTCTGTGGGGCCCCCTGTGGGCAAGATGGCATCCCTGGAGTCTAC 151
QY 814 ACCATATTGCAA 827
DB 150 ACCTATTATTGCAA 137

RESULT 3

BC035385

LOCUS

DEFINITION

Homo sapiens, similar to kallikrein 12, clone IMAGE:4825490, mRNA.

ACCESSION

BC035385

VERSION

BC035385.1

KEYWORDS

HTC.

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

1

REFERENCE 1 (bases 1 to 1186)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 71 Row: m Column: 9
This clone has the following problem: retained intron.

FEATURES
source
1..1186
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4825490"
/tissue_types="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN
Query Match 22.9%; Score 250; DB 11; Length 1186;
Best Local Similarity 100.0%; Pred. No. 1.8e-61;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 578 ACCATTCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTTCCATGCCACCTGCC 637
Db |||||
802 ACCATTCCGGATCTGCTCCAGTGCCTCAACCTCTCCATGCTTCCATGCCACCTGCC 861
QY 638 ATGCTGTGTATCCCGGAGAATCAAGAGCAACATGCTGTGCAAGCGGCGTCCCGGGC 697
Db |||||
862 ATGCTGTGTATCCCGGAGAATCAAGAGCAACATGCTGTGCAAGCGGCGTCCCGGGC 921
QY 698 AGGATGCTGCAGGGTGTCTGGGGCCCCCTGCTGTGTGTGGGGAGTCTTCAAGTTC 757
Db |||||
922 AGGATGCTGCAGGGTGTCTGGGGCCCCCTGCTGTGTGTGGGGAGTCTTCAAGTTC 981
QY 758 TGGTGTCTGGGGTCTGTGGGGCCCCCTGTGACAAAGATGGCATCCCTGGAGTCTACACCT 817
Db |||||
982 TGGTGTCTGGGGTCTGTGGGGCCCCCTGTGACAAAGATGGCATCCCTGGAGTCTACACCT 1041
QY 818 ATATTTGCAA 827
Db |||||
1042 ATATTTGCAA 1051

RESULT 4
BG230967
LOCUS
DEFINITION BG230967 264 bp mRNA linear EST 09-FEB-2001
similar to IR:Q3UK1-Q9UK1 XALLIKREIN-LIKE PROTEIN 5-RELATED
PROTEIN 1. ; mRNA sequence.
ACCESSION BG230967
VERSION BG230967.1 GI:12726030
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 264)
AUTHORS NCI/NIH-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D., Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 257.

FEATURES
source
1..264
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4227367"
/tissue_types="well-differentiated invasive carcinoma, floor of mouth"
/lab_host="DH10B"
/clone_lib="NCI CGAP HN8"
/note="Vector: PAMPI0; cDNA made by oligo-dT priming. Non-directionally cloned into the UDG sites of PAMPI0. Size-selected on agarose gel, average insert size 500 bp. Primary library; non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D (NCI). Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

ORIGIN
Query Match 19.8%; Score 216; DB 12; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.2e-51;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 857 GGAACTCACTGCTGTTTCTCCACCTCCACCCCACTTAACCTGGGTACCCCTCTG 916
Db |||||
43 GGAACTCACTGCTGTTTCTCCACCTCCACCCCACTTAACCTGGGTACCCCTCTG 102
QY 917 GCCTCAGACGACCAATATCTCTCCATCACTTCCCTAGCTCCACTCTTGTGGCTGG 976
Db |||||
103 GCCTCAGACGACCAATATCTCTCCATCACTTCCCTAGCTCCACTCTTGTGGCTGG 162
QY 977 GAACTTCTTGGAACTTTAACTCTCTGCGAGCCCTTTAAGACCCAGCGGGGTGAGAGA 1036
Db |||||
163 GAACTTCTTGGAACTTTAACTCTCTGCGAGCCCTTTAAGACCCAGCGGGGTGAGAGA 222
QY 1037 AGTGTGCAATAGTCTGGAATAAATAATGAAGGA 1072
Db |||||
223 AGTGTGCAATAGTCTGGAATAAATAATGAAGGA 258

RESULT 5
AI554072
LOCUS
DEFINITION AI554072 383 bp mRNA linear EST 13-APR-1999
te49g02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2090066 3', similar to TR:Q99920 Q99920 NES1=NORMAL EPITHELIAL
CELL SPECIFIC GENE 1. ; mRNA sequence.
ACCESSION AI554072
VERSION AI554072.1 GI:4486435
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE
AUTHORS      1 (bases 1 to 383)
TITLE        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
             National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-remail.nih.gov
             This clone is available royalty-free through LNLML; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert length: 575 Std Error: 0.00
             Seq primer: -40UP from Gibco
             High quality sequence stop: 373.
FEATURES
source
1..383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2090066"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match      9.3%; Score 101; DB 9; Length 383;
Best Local Similarity 100.0%; Pred. No. 5.9e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 CAGCCGTGCGAGTGGGGCTGTTTGAAGGCGACACAGCTGGCTGGGGGGTGCCTTATT 276
Db 192 CAGCCGTGCGAGTGGGGCTGTTTGAAGGCGACACAGCTGGCTGGGGGGTGCCTTATT 251
QY 277 GACCACAGTGGGTCTCTACAGCGGCTCACTGCAGCGGCAG 317
Db 252 GACCACAGTGGGTCTCTACAGCGGCTCACTGCAGCGGCAG 292
RESULT 6
BX103516/c
LOCUS          BX103516 Soares NFL T GBC S1 Homo sapiens cDNA clone
DEFINITION    IMAGE998D035152 ; IMAGE:2090066, mRNA sequence.
ACCESSION     BX103516
VERSION       BX103516
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 478)
AUTHORS       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (bases 1 to 478)
JOURNAL       Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
              Radelof, U., Schneider, D. and Korn, B.
COMMENT       Human Unigeneset - RZPD3
              Unpublished (2003)
              Contact: Ina Rofls
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
              RZPD; IMAGE998D035152.
              RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
              Human Unigeneset - RZPD3 (RZPDLIB No.972)
              http://www.rzpd.de/CloneCards/cgi-
              bin/showlib.pl.cgi?responsefileNo=972 Contact: Ina Rofls
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

```

```

Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTTACACAGGAACAGCTATGAC.
FEATURES
source
1..478
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998D035152 ; IMAGE:2090066"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI CGAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match      9.3%; Score 101; DB 13; Length 478;
Best Local Similarity 100.0%; Pred. No. 5e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 CAGCCGTGCGAGTGGGGCTGTTTGAAGGCGACACAGCTGGCTGGGGGGTGCCTTATT 276
Db 322 CAGCCGTGCGAGTGGGGCTGTTTGAAGGCGACACAGCTGGCTGGGGGGTGCCTTATT 263
QY 277 GACCACAGTGGGTCTCTACAGCGGCTCACTGCAGCGGCAG 317
Db 262 GACCACAGTGGGTCTCTACAGCGGCTCACTGCAGCGGCAG 222
RESULT 7
BG721150
LOCUS          BG721150 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4825490 5',
DEFINITION    mRNA sequence.
ACCESSION     BG721150
VERSION       BG721150
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 550)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
              cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
              Toshiyuki and Piero Carninci (RIKEN)
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM10738 row: 1 column: 03
              High quality sequence stop: 549.
FEATURES
source
1..550
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4825490"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (Gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN
Query Match          4.8%; Score 52; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 GGAAGTGACCCACCATGGGCTCAGCATCTTTTTCCTCTGTGTCTTGG 158
      |||||||
Db 436 GGAAGTGACCCACCATGGGCTCAGCATCTTTTTCCTCTGTGTCTTGG 487

RESULT 8
BQ129094/c
LOCUS
DEFINITION
BQ129094 577 bp mRNA linear EST 19-APR-2002
i127h02.x1 Melton Amplified Mouse E16 5 Pancreas 3 M16S1 A Mus
musculus cDNA clone IMAGE:5943026 3' similar to TR:Q9Z1H1 Q9Z1H1
P6-5 GENE, 3' END ;, mRNA sequence.

ACCESSION
BQ129094
VERSION
BQ129094.1 GI:20203005
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 577)
AUTHORS
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center This clone is
available royalty-free through LLNL; please contact the IMAGE
consortium (info@image.llnl.gov) for further information
Seq primer: -40UP from Gibco
High quality sequence stop: 235.
FEATURES
Location/Qualifiers
source
1..577
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5943026"
/sex="Both"
/tissue_type="pancreas"
/dev_stage="Embryonic day E16.5"
/lab_host="DH10B"
/clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 M16S1

```

```

A"
/notes="Vector: pSPORT1; Site 1: Not 1; Site 2: Sal 1;
Library constructed using SuperScript Plasmid Library kit
(Life Technologies). cDNA made by oligo-dT priming.
Size-selected by column fractionation; average insert size
0.97 kb. Amplified once on solid support. cDNA Library
Preparation: Guolin Chen."

ORIGIN
Query Match          2.4%; Score 26; DB 13; Length 577;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 TGCACGGGTGATTCCTGGGGCCCCCT 731
      |||||||
Db 463 TGCACGGGTGATTCCTGGGGCCCCCT 438

RESULT 9
BQ270082
LOCUS
DEFINITION
BQ270082 600 bp mRNA linear EST 15-JUL-2003
ik3302.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5782850 5'
similar to SW:TRY2_HUMAN P07478 TRYPSINOGEN II PRECURSOR ;, mRNA
sequence.

ACCESSION
BQ270082
VERSION
BQ270082.1 GI:20495148
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 600)
AUTHORS
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 422.
FEATURES
Location/Qualifiers
source
1..600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5782850"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/notes="Organ: Pancreas;
Vector: pBluescript SK(-); Site 1:
NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

```



```

Query Match      2.3%; Score 25; DB 13; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 704 CTGCGAGGGTGATTCTGGGGGCC 728
|||||
566 CTGCGAGGGTGATTCTGGGGGCC 590
|||||

Db
RESULT 10
LOCUS BQ269009
DEFINITION ik19g10.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5781690 5',
similar to SW:TRY2_HUMAN P07478 TRYPSINOGEN II PRECURSOR ;, mRNA
sequence.
ACCESSION BQ269009
VERSION BQ269009.1 GI:20494075
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 636)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other ESTs: ik19g10.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 478.
Location/Qualifiers
1. .636
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5781690"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN
Query Match      2.3%; Score 25; DB 13; Length 636;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 704 CTGCGAGGGTGATTCTGGGGGCC 728
|||||
566 CTGCGAGGGTGATTCTGGGGGCC 590
|||||

Db
590 CTGCGAGGGTGATTCTGGGGGCC 614

RESULT 11
LOCUS BI794700/c
DEFINITION i63e04.x2 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
musculus cDNA clone IMAGE:5658342 3', similar to SW:EUL_PIG P00772
ELASTASE 1 PRECURSOR ;, mRNA sequence.
ACCESSION BI794700
VERSION BI794700.1 GI:15822425
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 283)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other ESTs: ic63e04.y2
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Trace considered overall poor quality
MGI:194468 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 283
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5658342"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stages="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
NI-MMS1"
/note="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
was made by oligo-dT priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram
single-stranded mixed library plasmid DNA was mixed with
5 micrograms PCR product representing mixed library
inserts and hybridized to an EcoT of 6. Single-stranded
(unhybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."

```

ORIGIN

Query Match 2.2%; Score 24; DB 12; Length 283;
 Best Local Similarity 100.0%; Pred. No. 4.5e+03;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 708 CCAGGCTGATCTCGGGGCCCCCT 731
 Db 167 CCAGGCTGATCTCGGGGCCCCCT 144

RESULT 12

AI322408

LOCUS

DEFINITION AI322408 449 bp mRNA linear EST 23-DEC-1998
 IMAGE:464085 5' similar to SW:TRY3_SALSA P35033 TRYPSINOGEN III
 PRECURSOR ; mRNA sequence.

ACCESSION

AI322408

VERSION

AI322408.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MG1:277901

This read is a RESEQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Putative full length read
 vector to vector length is 607
 Seq primer: -40Rp from Gibco
 High quality sequence stop: 408.

FEATURES

source

1. .449

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:464085"

/dev_stage="19.5 dpc total fetus"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares mouse p3MW19.5"

/note="Vector: p773D (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGGCGGCATTTTTTTTTTTTTTTT 3']

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified p773 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento

Soares and M.Patima Bonaldo. RNA was kindly provided by

Dr. Minoru Ko (Wayne State University)."

Matches

23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 594 GTCCAGTGCTCAACCTCTCCA 616
 Db 147 GTCCAGTGCTCAACCTCTCCA 169

RESULT 13

CA313295/c

LOCUS

DEFINITION CA313295 451 bp mRNA linear EST 04-NOV-2002
 UI-CF-FNO-aeY-o-05-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
 UI-CF-FNO-aeY-o-05-0-UI 3', mRNA sequence.

ACCESSION

CA313295

VERSION

CA313295.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 154-233, >(GAA)n#Simple_repeat
 Seq primer: M13 FORWARD
 POLYA=Yes.

Location/Qualifiers
 1. .451
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-FNO-aeY-o-05-0-UI"
 /tissue_type="Human Lung Epithelial cells"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-FNO"
 /note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-FNO is a subtracted cDNA library derived from two
 normalized Human lung epithelial cell libraries (EN1 and
 DUL) the library was subtracted according to according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. For additional information, contact:
 bento-soares@uiowa.edu
 TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
 TAG LIB=UI-CF-FNO
 TAG_SEQ=GGCTGTAGGC"

FEATURES

source

1. .451

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-FNO-aeY-o-05-0-UI"

/tissue_type="Human Lung Epithelial cells"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-FNO"

/note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-FNO is a subtracted cDNA library derived from two
 normalized Human lung epithelial cell libraries (EN1 and
 DUL) the library was subtracted according to according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. For additional information, contact:
 bento-soares@uiowa.edu
 TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
 TAG LIB=UI-CF-FNO
 TAG_SEQ=GGCTGTAGGC"

ORIGIN

Query Match 2.1%; Score 23; DB 14; Length 451;
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1069 AGGAGGGCAAAAAAAAAAAAAA 1091
 Db 24 AGGAGGGCAAAAAAAAAAAAAA 2

RESULT 14
AA028356
LOCUS
DEFINITION
m120d11.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:464085 5' similar to SW:TRIP_BOVIN P00760 TRYPSINOGEN ; mRNA
sequence.

AA028356
AA028356.1 GI:1494434
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 506)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:277901

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 486.

FEATURES
source
1..506
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:464085"
/dev_stage="19.5 dpc total fetus"
/lab_host="PH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: pT73D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACCAATCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

ORIGIN
Query Match 2.1%; Score 23; DB 9; Length 506;
Best Local Similarity 100.0%; Pred. No. 5.7e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 GCTCCAGTGCCTCAACCTCTCCA 616
Db 146 GCTCCAGTGCCTCAACCTCTCCA 168

RESULT 15
BX514743
LOCUS
DEFINITION
BX514743 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:464085 5' similar to SW:TRIP_BOVIN P00760 TRYPSINOGEN ; mRNA
sequence.
ACCESSION
VERSION
BX514743.1 GI:32244696

KEYWORDS
SOURCE
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 520)
Heil, O., Ebert, L., Neubert, P., Peters, M., Radelof, U., Schneider, D.
and Korn, B.
TITLE
Mouse Unigeneset - RZPD2
JOURNAL
COMMENT
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGP998N221098
RZFDLIB; I.M.A.G.E. cDNA Clone Collection;
Mouse Unigeneset - RZPD2 (RZFDLIB No.981)
http://www.rzpd.de/Clonecards/cgi-
bin/showlib.pl.cgi/response?libNo=981 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 5, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
T7, Primer sequence: TAATACGACTCTACTATAGGG.

FEATURES
source
1..520
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGP998N221098 ; IMAGE:464085"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: pT73D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACCAATCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

ORIGIN
Query Match 2.1%; Score 23; DB 13; Length 520;
Best Local Similarity 100.0%; Pred. No. 5.6e+03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 GCTCCAGTGCCTCAACCTCTCCA 616
Db 148 GCTCCAGTGCCTCAACCTCTCCA 170

Search completed: June 20, 2004, 09:46:47
Job time : 3358 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:26:06 ; Search time 60 Seconds
(without alignments)
1167.863 Million cell updates/sec

Title: US-10-015-385A-194

Perfect score: 1374
Sequence: 1 MGLSIFLLCVLGLSQATP.....GVYYICKYVDIRMIMRN 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1374	100.0	248	3 AAB21304	Aab21304 Human KIK
2	1374	100.0	248	3 AAB24428	Aab24428 Human PRO
3	1374	100.0	248	3 AAB24032	Aab24032 Human PRO
4	1374	100.0	248	3 AAY99393	Aay99393 Human PRO
5	1374	100.0	248	4 AAM23994	Aam23994 Human EST
6	1374	100.0	248	4 AAB66142	Aab66142 Protein o
7	1374	100.0	248	6 ABO33635	Abc33635 Novel hum
8	1374	100.0	248	7 ABO44488	Abc44488 Human sec
9	1374	100.0	248	7 ABO33512	Abc33512 Novel hum
10	1374	100.0	248	7 ADC18063	Adc18063 Human PRO
11	1374	100.0	248	7 ADD70709	Add70709 Human sec
12	1374	100.0	248	7 ADD39786	Add39786 Human sec
13	1374	100.0	248	7 ADD70232	Add70232 Human sec
14	1374	100.0	248	7 ADD38353	Add38353 Human sec
15	1374	100.0	248	7 ADD39309	Add39309 Human sec
16	1374	100.0	248	7 ADD38832	Add38832 Human sec
17	1374	100.0	248	7 ADD40263	Add40263 Human sec
18	1374	100.0	248	7 ADE50484	Ade50484 Human sec
19	1374	100.0	248	7 ADE20096	Ade20096 Human sec
20	1374	100.0	248	7 ADE50007	Ade50007 Human sec
21	1374	100.0	248	7 ADE21565	Ade21565 Human sec
22	1367	99.5	248	6 AAO29516	Aao29516 Human kal
23	1301	94.7	254	3 AAB21303	Aab21303 Human KIK
24	1301	94.7	254	5 ABG66676	Abg66676 Human nov
25	1019	74.2	184	3 AAB21301	Aab21301 Human KIK

ALIGNMENTS

RESULT 1

ID AAB21304 standard; protein; 248 AA.

XX AAB21304;

DT 02-FEB-2001 (first entry)

XX Human KIK-L5 protein #4.

Human; KIK-L1; KIK-L2; KIK-L3; KIK-L4; KIK-L5; KIK-L6;
kallikrein-like protein; serine protease; cytostatic; cancer;
prostate cancer.

OS Homo sapiens.

XX WO200053776-A2.

PD 14-SEP-2000.

XX 09-MAR-2000; 2000WO-CA000258.

XX 11-MAR-1999; 99US-0124260P.

XX 01-APR-1999; 99US-0127386P.

XX 21-JUL-1999; 99US-0144919P.

XX (MOUN) MOUNT SINAI HOSPITAL.

XX Yousef GM, Diamandis EP;

XX WPI; 2000-587440/55.

XX N-PSDB; AAA95944.

XX New kallikrein-like (KIK-L) proteins for diagnosing and treating KIK-L protein mediated disorders, especially cancer.
Claim 12; Page 172; 184pp; English.

The present sequence is one of four alternatively spliced kallikrein-like proteins encoded by the human KIK-L4 gene. Kallikreins and kallikrein-like proteins are a subgroup of the serine protease enzyme family. They catalyse the selective cleavage of specific polypeptide precursors to release peptides with potent biological activity. Nucleic acids encoding kallikrein-like proteins KIK-L1, KIK-L2, KIK-L3, KIK-L4, KIK-L5 and KIK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used

CC to treat conditions mediated by the kallikrein-like proteins

XX Sequence 248 AA;

Query Match 100.0%; Score 1374; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGVLIDHRWLTA 60
QY 61 AHCSGRYVWRLGHSLSQDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
DB 61 AHCSGRYVWRLGHSLSQDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
QY 121 TSSVQPLPLPNDCAATGTECHVSGWGIINHPNPPDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCAATGTECHVSGWGIINHPNPPDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDACCQDGGPLVCGVGLQGLVSWGSGVPGCGQDIPGVYTIKYYVDW 240
DB 181 TSNMVCAGVPGQDACCQDGGPLVCGVGLQGLVSWGSGVPGCGQDIPGVYTIKYYVDW 240
QY 241 IRMIMRNN 248
DB 241 IRMIMRNN 248

RESULT 2

AAB24428
ID AAB24428 standard; protein; 248 AA.

XX AAB24428;

DT 07-NOV-2000 (first entry)
XX

DE Human PRO1303 protein sequence SEQ ID NO:203.

XX Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW angiogenic; proliferative; cardiast; cardiovascular; antiatherosclerotic;
KW cytostatic; gene therapy; vaccine.

XX Homo sapiens.

XX W0200032221-A2.

XX 08-JUN-2000.

XX 30-NOV-1999; 99WO-US028313.

XX 01-DEC-1998; 98WO-US025108.

XX 16-DEC-1998; 98US-0112850P.

XX 12-JAN-1999; 99US-0115554P.

XX 08-MAR-1999; 99WO-US005028.

XX 12-MAR-1999; 99US-0123357P.

XX 28-APR-1999; 99US-0131445P.

XX 14-MAY-1999; 99US-0134287P.

XX 23-JUN-1999; 99WO-US012252.

XX 20-JUL-1999; 99US-0141037P.

XX 26-JUL-1999; 99US-0144758P.

XX 01-SEP-1999; 99WO-US020111.

XX 08-SEP-1999; 99WO-US020594.

XX 13-SEP-1999; 99WO-US020944.

XX 15-SEP-1999; 99WO-US021090.

XX 05-OCT-1999; 99WO-US021547.

XX 29-OCT-1999; 99WO-US023089.

XX 29-OCT-1999; 99US-0162506P.

XX (GETH) GENENTECH INC.

XX PA

XX

PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ;
PI Goddard A, Godowski FU, Gurney AL, Klein RD, Kuo SS, Paoni NF;
PI Smith V, Watanabe CK, Williams PM, Wood WI;

XX WPI; 2000-412154/35.
XX N-PSDB; AAB77671.

XX Nucleic acids encoding PRO polypeptides useful for preventing, diagnosing
PT and treating diagnosing a cardiovascular, endothelial or angiogenic
PT disorders in mammals.

XX Claim 72; Fig 82; 315pp; English.

XX The present invention describes nucleic acids encoding PRO polypeptides
XX useful for preventing, diagnosing and treating diagnosing a
XX cardiovascular, endothelial or angiogenic disorder in mammals by
XX modulating cell proliferation, angiogenesis and cardiovascularisation,
XX and for identifying agonists and antagonists of these processes. The
XX nucleic acids and the proteins they encode may be used in the prevention,
XX treatment and diagnosis of diseases associated with inappropriate PRO
XX expression such as cardiovascular, endothelial or angiogenic disorders in
XX mammals (e.g. atherosclerosis, cancers and cardiac hypertrophy). For
XX example, the nucleic acids (NCs) and vectors containing them and the PRO
XX polypeptide may be used to treat disorders associated with decreased PRO
XX expression. AAB77510 to AAB77721 and AAB24388 to AAB24435 represent
XX nucleotide and protein sequences used in the exemplification of the
XX present invention

XX Sequence 248 AA;

Query Match 100.0%; Score 1374; DB 3; Length 248;

Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGVLIDHRWLTA 60

QY 61 AHCSGRYVWRLGHSLSQDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120
DB 61 AHCSGRYVWRLGHSLSQDWTQIRHSGFSVTHPGVLGASTSHEHDLRLRLPLVRV 120

QY 121 TSSVQPLPLPNDCAATGTECHVSGWGIINHPNPPDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCAATGTECHVSGWGIINHPNPPDLLQCLNLSIVSHATCHGVYPGRI 180

QY 181 TSNMVCAGVPGQDACCQDGGPLVCGVGLQGLVSWGSGVPGCGQDIPGVYTIKYYVDW 240
DB 181 TSNMVCAGVPGQDACCQDGGPLVCGVGLQGLVSWGSGVPGCGQDIPGVYTIKYYVDW 240

QY 241 IRMIMRNN 248
DB 241 IRMIMRNN 248

RESULT 3

AAB24032
ID AAB24032 standard; protein; 248 AA.

XX AAB24032;

XX 25-JAN-2001 (first entry)

XX Human PRO1303 protein sequence SEQ ID NO:33.

XX Human; tumour; diagnosis; neoplastic disease; proliferation; cancer;
KW identification; tumourigenesis; anticancer; detection.

XX Homo sapiens.

XX W0200053750-A1.

XX 14-SEP-2000.

XX PF 02-DEC-1999; 99WO-US028551.
XX PR 08-MAR-1999; 99WO-US005028.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 29-OCT-1999; 99US-0162506P.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 01-DEC-1999; 99WO-US028634.
XX PA (GETH) GENENTECH INC.
XX PI Botstein D, Goddard A, Gurney AL, Roy MA, Watanabe CX, Wood WI;
XX DR WPI; 2000-594320/56.
XX DR N-PSDB; AAC58114.
XX PT Antibodies specific for PRO polypeptides, used to diagnose and inhibit
XX PT the growth of tumors in mammals, and to identify inhibitors of PRO
XX PT polypeptide activity or expression.
XX PS Claim 61; Fig 24; 226pp; English.
XX CC The present invention describes an antibody that binds to a human protein
XX CC (I) selected from: PRO381; PRO1269; PRO1410; PRO1755; PRO1780; PRO3434;
XX CC PRO1927; PRO3567; PRO1295; PRO1293; PRO1303; PRO4344; PRO4397;
XX CC PRO4407; PRO1555; PRO1096; PRO2038; and PRO2262. (I) has anticancer
XX CC activity and can be used to diagnose tumours in mammals, by detecting
XX CC complex formation when the antibody is contacted with test cells.
XX CC Increased expression of genes encoding (I) can also be detected to
XX CC diagnose tumours. Agents which inhibit the activity of (I), especially
XX CC the antibodies, or an antisense oligonucleotide which hybridises to genes
XX CC encoding (I), can be used to inhibit tumour growth, preferably by
XX CC inducing cell death. Methods from the present invention can be used to
XX CC identify compounds which inhibit the biological activity of (I). AAC58019
XX CC to AAC58102 represent PCR primers and hybridisation probes used in
XX CC examples from the present invention for human PRO sequences. AAC58103 to
XX CC AAC58122 and AAB24021 to AAB24040 represent human PRO polynucleotide and
XX CC protein sequences given in the exemplification of the present invention
XX SQ Sequence 248 AA;

Query Match 100.0%; Score 1374; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLICVLGSLQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGVLIDHRWLTA 60
DB 1 MGLSIFLLICVLGSLQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGVLIDHRWLTA 60

QY 61 AHCSSRYWVRIGEHSLQDLTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120
DB 61 AHCSSRYWVRIGEHSLQDLTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120

QY 121 TSSVQPLPNDCAATAGTECHVSGWGITNHRNPPDQLQCLNLSIVSHATGCHVYPGRI 180
DB 121 TSSVQPLPNDCAATAGTECHVSGWGITNHRNPPDQLQCLNLSIVSHATGCHVYPGRI 180

QY 181 TSNMVCAGVPQDAGCGSGGGLVCGGVLQGLVSGVSPGCGDGPVGYVYCKYVDW 240
DB 181 TSNMVCAGVPQDAGCGSGGGLVCGGVLQGLVSGVSPGCGDGPVGYVYCKYVDW 240

QY 241 IRMIMRNN 248
DB 241 IRMIMRNN 248

RESULT 4
AAY99393
ID AAY99393 standard; protein; 248 AA.
XX AC AAY99393;
XX AC AAY99393;
XX DB 08-AUG-2000 (first entry)

XX DE Human PRO1303 (UNQ669) amino acid sequence SEQ ID NO:194.
XX KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
XX KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
XX OS Homo sapiens.
XX PN WO200012708-A2.
XX PD 09-MAR-2000.
XX PF 01-SEP-1999; 99WO-US020111.
XX PR 01-SEP-1998; 98US-0098716P.
XX PR 01-SEP-1998; 98US-0098749P.
XX PR 02-SEP-1998; 98US-0098750P.
XX PR 02-SEP-1998; 98US-0098803P.
XX PR 02-SEP-1998; 98US-0098821P.
XX PR 02-SEP-1998; 98US-0098843P.
XX PR 02-SEP-1998; 98US-0098536P.
XX PR 02-SEP-1998; 98US-0098596P.
XX PR 02-SEP-1998; 98US-0098598P.
XX PR 02-SEP-1998; 98US-0098602P.
XX PR 02-SEP-1998; 98US-0098642P.
XX PR 02-SEP-1998; 98US-0098741P.
XX PR 10-SEP-1998; 98US-0098754P.
XX PR 10-SEP-1998; 98US-0098763P.
XX PR 10-SEP-1998; 98US-0098792P.
XX PR 10-SEP-1998; 98US-0098808P.
XX PR 10-SEP-1998; 98US-0098812P.
XX PR 10-SEP-1998; 98US-0098815P.
XX PR 10-SEP-1998; 98US-0098816P.
XX PR 15-SEP-1998; 98US-0100385P.
XX PR 15-SEP-1998; 98US-0100388P.
XX PR 15-SEP-1998; 98US-0100390P.
XX PR 15-SEP-1998; 98US-0100584P.
XX PR 16-SEP-1998; 98US-0100627P.
XX PR 16-SEP-1998; 98US-0100661P.
XX PR 16-SEP-1998; 98US-0100662P.
XX PR 16-SEP-1998; 98US-0100664P.
XX PR 16-SEP-1998; 98US-0100683P.
XX PR 17-SEP-1998; 98US-0100684P.
XX PR 17-SEP-1998; 98US-0100710P.
XX PR 17-SEP-1998; 98US-0100711P.
XX PR 17-SEP-1998; 98US-0100919P.
XX PR 17-SEP-1998; 98US-0100930P.
XX PR 18-SEP-1998; 98US-0100849P.
XX PR 18-SEP-1998; 98US-0100849P.
XX PR 18-SEP-1998; 98US-0101049P.
XX PR 18-SEP-1998; 98US-0101068P.
XX PR 18-SEP-1998; 98US-0101071P.
XX PR 22-SEP-1998; 98US-0101279P.
XX PR 23-SEP-1998; 98US-0101471P.
XX PR 23-SEP-1998; 98US-0101472P.
XX PR 23-SEP-1998; 98US-0101473P.
XX PR 23-SEP-1998; 98US-0101474P.
XX PR 23-SEP-1998; 98US-0101475P.
XX PR 23-SEP-1998; 98US-0101476P.
XX PR 23-SEP-1998; 98US-0101477P.
XX PR 24-SEP-1998; 98US-0101479P.
XX PR 24-SEP-1998; 98US-0101738P.
XX PR 24-SEP-1998; 98US-0101741P.
XX PR 24-SEP-1998; 98US-0101743P.
XX PR 24-SEP-1998; 98US-0101915P.
XX PR 24-SEP-1998; 98US-0101916P.
XX PR 24-SEP-1998; 98US-0102207P.
XX PR 29-SEP-1998; 98US-0102240P.
XX PR 29-SEP-1998; 98US-0102307P.
XX PR 29-SEP-1998; 98US-0102330P.
XX PR 29-SEP-1998; 98US-0102331P.
XX PR 30-SEP-1998; 98US-0102484P.
XX PR 30-SEP-1998; 98US-0102487P.
XX PR 30-SEP-1998; 98US-0102570P.


```
XX WPI; 2001-476164/51.
DR N-PSDB; AAH98653.
XX
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use.
XX
XX Claim 20; Page 1048-1049; 1275pp; English.
XX
XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 1374; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRGCGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRGCGVLIDHRWLTA 60
QY 61 AHCSGSRVWVRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLRV 120
DB 61 AHCSGSRVWVRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPFDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPFDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDACQDGGPLVCGVGLQGLVSGVSGPCGQDGPVVTYICKYVDW 240
DB 181 TSNMVCAGVPGQDACQDGGPLVCGVGLQGLVSGVSGPCGQDGPVVTYICKYVDW 240
QY 241 IRMTMRNN 248
DB 241 IRMTMRNN 248
RESULT 6
AAB66142
ID AAB66142 standard; protein; 248 AA.
XX
AC AAB66142;
XX
DT 02-APR-2001 (first entry)
XX
DE Protein of the invention #54.
XX
KW Secreted; transmembrane; gene therapy.
XX
OS Unidentified.
XX
PN W0200078961-A1.
XX
XX 28-DEC-2000.
XX
XX 18-FEB-2000; 2000WO-US004342.
XX
XX 23-JUN-1999; 99US-0141037P.
XX
XX 20-JUL-1999; 99US-0144758P.
XX
XX 26-JUL-1999; 99US-0145698P.
XX
XX 01-SEP-1999; 99WO-US020111.
XX
XX 29-OCT-1999; 99US-0162506P.
XX
XX 30-NOV-1999; 99WO-US028313.
XX
XX 02-DEC-1999; 99WO-US028551.
XX
XX 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,
PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tamas D, Watanabe CK,
PI Williams PM, Wood WI;
XX
XX WPI; 2001-071395/08.
XX
XX Secreted and transmembrane proteins and nucleic acids designated PRO,
PT useful as hybridization probes, in chromosome and gene mapping and gene
PT therapy.
XX
XX Claim 1; Fig 108; 787pp; English.
XX
XX The present invention relates to secreted and transmembrane proteins.
CC These proteins and the DNA encoding them may be used as hybridization
CC probes, in chromosome and gene mapping and in the generation of anti-
CC sense RNA and DNA. They may also be used to generate either
CC transgenic animals or knockout animals which are in turn useful for
CC development and screening of therapeutically useful reagents. The nucleic
CC acids may also be used in gene therapy
XX
SQ Sequence 248 AA;
Query Match 100.0%; Score 1374; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRGCGVLIDHRWLTA 60
DB 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRGCGVLIDHRWLTA 60
QY 61 AHCSGSRVWVRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLRV 120
DB 61 AHCSGSRVWVRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLRV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPFDLLQCLNLSIVSHATCHGVYPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNRPFDLLQCLNLSIVSHATCHGVYPGRI 180
QY 181 TSNMVCAGVPGQDACQDGGPLVCGVGLQGLVSGVSGPCGQDGPVVTYICKYVDW 240
DB 181 TSNMVCAGVPGQDACQDGGPLVCGVGLQGLVSGVSGPCGQDGPVVTYICKYVDW 240
QY 241 IRMTMRNN 248
DB 241 IRMTMRNN 248
RESULT 7
ABO33635
ID ABO33635 standard; protein; 248 AA.
XX
AC ABO33635;
XX
DT 17-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1303.
XX
XX Human; secreted and transmembrane protein; PRO; angiogenesis;
XX endothelial cell proliferation; wound healing; immune response;
XX T-lymphocytes proliferation; neonatal heart hypertrophy; tumour;
XX cardiac insufficiency disorder; calcium flux; inflammation;
XX vascular endothelial growth factor-stimulated proliferation;
XX mammalian kidney mesangial cell proliferation; Berger disease;
XX nephropathy; Schanlein-Henoch purpura; celiac disease; Crohn's disease;
XX dermatitis herpetiformis; diabetes; haemoglobin switch; insulinemia;
XX pancreatic beta-cell precursor cell differentiation; thalassemia;
XX obesity; auditory hair cell regeneration; hearing loss; bone disorder;
```

KW cartilage disorder; sports injury; arthritis.
XX Homo sapiens.
OS US2003073130-A1.
PN 17-APR-2003.
XX
XX 11-DEC-2001; 2001US-00015863.
XX
XX 01-SEP-1998; 98US-0098716P.
XX 01-SEP-1998; 98US-0098723P.
XX 01-SEP-1998; 98US-0098749P.
XX 01-SEP-1998; 98US-0098750P.
XX 02-SEP-1998; 98US-0098803P.
XX 02-SEP-1998; 98US-0098821P.
XX 02-SEP-1998; 98US-0098843P.
XX 02-SEP-1998; 98US-0098536P.
XX 09-SEP-1998; 98US-0099596P.
XX 09-SEP-1998; 98US-0099598P.
XX 09-SEP-1998; 98US-0099602P.
XX 09-SEP-1998; 98US-0099642P.
XX 10-SEP-1998; 98US-0099741P.
XX 10-SEP-1998; 98US-0099754P.
XX 10-SEP-1998; 98US-0099763P.
XX 10-SEP-1998; 98US-0099792P.
XX 10-SEP-1998; 98US-0099808P.
XX 10-SEP-1998; 98US-0099812P.
XX 10-SEP-1998; 98US-0099815P.
XX 15-SEP-1998; 98US-0099816P.
XX 15-SEP-1998; 98US-0100385P.
XX 15-SEP-1998; 98US-0100388P.
XX 15-SEP-1998; 98US-0100390P.
XX 16-SEP-1998; 98US-0100584P.
XX 16-SEP-1998; 98US-0100627P.
XX 16-SEP-1998; 98US-0100661P.
XX 16-SEP-1998; 98US-0100662P.
XX 16-SEP-1998; 98US-0100664P.
XX 17-SEP-1998; 98US-0100683P.
XX 17-SEP-1998; 98US-0100684P.
XX 17-SEP-1998; 98US-0100710P.
XX 17-SEP-1998; 98US-0100711P.
XX 17-SEP-1998; 98US-0100919P.
XX 17-SEP-1998; 98US-0100930P.
XX 18-SEP-1998; 98US-0100848P.
XX 18-SEP-1998; 98US-0100849P.
XX 18-SEP-1998; 98US-0101014P.
XX 18-SEP-1998; 98US-0101071P.
XX 18-SEP-1998; 98US-0101076P.
XX 18-SEP-1998; 98US-0101476P.
XX 23-SEP-1998; 98US-0101477P.
XX 23-SEP-1998; 98US-0101479P.
XX 23-SEP-1998; 98US-0101471P.
XX 23-SEP-1998; 98US-0101472P.
XX 23-SEP-1998; 98US-0101474P.
XX 23-SEP-1998; 98US-0101475P.
XX 23-SEP-1998; 98US-0101475P.
XX 23-SEP-1998; 98US-0101476P.
XX 23-SEP-1998; 98US-0101477P.
XX 23-SEP-1998; 98US-0101479P.
XX 24-SEP-1998; 98US-0101738P.
XX 24-SEP-1998; 98US-0101741P.
XX 24-SEP-1998; 98US-0101741P.
XX 24-SEP-1998; 98US-0101915P.
XX 24-SEP-1998; 98US-0101916P.
XX 29-SEP-1998; 98US-0102070P.
XX 29-SEP-1998; 98US-0102240P.
XX 29-SEP-1998; 98US-0102307P.
XX 29-SEP-1998; 98US-0102330P.
XX 29-SEP-1998; 98US-0102331P.
XX 30-SEP-1998; 98US-0102484P.
XX 30-SEP-1998; 98US-0102570P.
XX 30-SEP-1998; 98US-0102570P.
XX 30-SEP-1998; 98US-0102571P.
XX 01-OCT-1998; 98US-0102684P.
XX 01-OCT-1998; 98US-0102687P.
XX 02-OCT-1998; 98US-0102965P.
XX 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105812P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 05-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 26-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.

29-SEP-1998; 98US-0102331P.
 30-SEP-1998; 98US-0102484P.
 30-SEP-1998; 98US-0102487P.
 30-SEP-1998; 98US-0102570P.
 30-SEP-1998; 98US-0102571P.
 01-OCT-1998; 98US-0102684P.
 01-OCT-1998; 98US-0102687P.
 02-OCT-1998; 98US-0102965P.
 06-OCT-1998; 98US-0103258P.
 06-OCT-1998; 98US-0103449P.
 07-OCT-1998; 98US-0103314P.
 07-OCT-1998; 98US-0103315P.
 07-OCT-1998; 98US-0103328P.
 07-OCT-1998; 98US-0103395P.
 07-OCT-1998; 98US-0103396P.
 07-OCT-1998; 98US-0103401P.
 08-OCT-1998; 98US-0103633P.
 08-OCT-1998; 98US-0103678P.
 08-OCT-1998; 98US-0103711P.
 14-OCT-1998; 98US-0104257P.
 20-OCT-1998; 98US-0104987P.
 20-OCT-1998; 98US-0105000P.
 20-OCT-1998; 98US-0105002P.
 21-OCT-1998; 98US-0105104P.
 22-OCT-1998; 98US-0105169P.
 22-OCT-1998; 98US-0105266P.
 26-OCT-1998; 98US-0105693P.
 26-OCT-1998; 98US-0105694P.
 27-OCT-1998; 98US-0105807P.
 27-OCT-1998; 98US-0105881P.
 27-OCT-1998; 98US-0105882P.
 27-OCT-1998; 98US-0106062P.
 28-OCT-1998; 98US-0106023P.
 28-OCT-1998; 98US-0106029P.
 28-OCT-1998; 98US-0106030P.
 28-OCT-1998; 98US-0106032P.
 28-OCT-1998; 98US-0106033P.
 28-OCT-1998; 98US-0106178P.
 29-OCT-1998; 98US-0106248P.
 29-OCT-1998; 98US-0106384P.
 30-OCT-1998; 98US-0108500P.
 30-OCT-1998; 98US-0108464P.
 03-NOV-1998; 98US-0106856P.
 03-NOV-1998; 98US-0106902P.
 03-NOV-1998; 98US-0106905P.
 03-NOV-1998; 98US-0106919P.
 03-NOV-1998; 98US-0106932P.
 03-NOV-1998; 98US-0106934P.
 10-NOV-1998; 98US-0107783P.
 17-NOV-1998; 98US-0108775P.
 17-NOV-1998; 98US-0108779P.
 17-NOV-1998; 98US-0108787P.
 17-NOV-1998; 98US-0108788P.
 17-NOV-1998; 98US-0108801P.
 17-NOV-1998; 98US-0108802P.
 17-NOV-1998; 98US-0108806P.
 17-NOV-1998; 98US-0108807P.
 17-NOV-1998; 98US-0108867P.
 17-NOV-1998; 98US-0108925P.
 18-NOV-1998; 98US-0108848P.
 18-NOV-1998; 98US-0108849P.
 18-NOV-1998; 98US-0108850P.
 18-NOV-1998; 98US-0108851P.
 18-NOV-1998; 98US-0108852P.
 18-NOV-1998; 98US-0108858P.
 18-NOV-1998; 98US-0108904P.
 22-DEC-1998; 98US-0113296P.
 30-DEC-1998; 98US-0114223P.
 05-JAN-1999; 99WO-US000106.
 16-APR-1999; 99US-0129674P.
 23-JUN-1999; 99US-0141037P.
 20-JUL-1999; 99US-0144758P.

26-JUL-1999; 99US-0145698P.
 01-SEP-1999; 99WO-US020111.
 15-SEP-1999; 99WO-US021194.
 29-OCT-1999; 99US-0162506P.
 30-NOV-1999; 99WO-US028313.
 02-DEC-1999; 99WO-US028551.
 16-DEC-1999; 99WO-US030095.
 05-JAN-2000; 2000WO-US000219.
 11-FEB-2000; 2000WO-US000376.
 18-FEB-2000; 2000WO-US003565.
 24-FEB-2000; 2000WO-US004342.
 02-MAR-2000; 2000WO-US005004.
 12-MAR-2000; 2000WO-US005841.
 15-MAR-2000; 2000WO-US006884.
 17-MAY-2000; 2000WO-US013705.
 22-MAY-2000; 2000WO-US014042.
 30-MAY-2000; 2000WO-US014941.
 02-JUN-2000; 2000WO-US015264.
 23-AUG-2000; 2000WO-US023522.
 24-AUG-2000; 2000WO-US023328.
 08-NOV-2000; 2000WO-US030952.
 10-NOV-2000; 2000WO-US030873.
 01-DEC-2000; 2000WO-US032678.
 28-FEB-2001; 2001WO-US006520.
 01-MAR-2001; 2001WO-US006666.
 01-JUN-2001; 2001WO-US017800.
 20-JUN-2001; 2001WO-US019692.
 29-JUN-2001; 2001WO-US021066.
 09-JUL-2001; 2001WO-US021735.
 04-SEP-2001; 2001US-00946374.
 XX (GETH) GENENTECH INC.
 PA Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
 PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tamas D, Watanabe CK;
 PI Williams PM, Wood WI;
 XX WPI; 2003-492259/46.
 DR N-P8DB; ACH04480.
 XX Novel secreted and transmembrane polypeptides and polynucleotides
 PT encoding them useful for treating various cardiac insufficiency
 PT disorders, bone and/or cartilage disorders such as sports injuries and
 PT arthritis.
 Query Match 100.0%; Score 1374; DB 7; Length 248;
 Best Local Similarity 100.0%; Pred. No. 3.8e-98;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGLSIFLLACVLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60
 DB 1 MGLSIFLLACVLGSLQAATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWLTA 60
 QY 61 AHCSSGRYVWRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120
 DB 61 AHCSSGRYVWRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120
 QY 121 TSSVQPLPLPNDCAATAGTECHVSGWGITNHPNRPDILLQCLNLSIVSHATGCHVPGRI 180
 DB 121 TSSVQPLPLPNDCAATAGTECHVSGWGITNHPNRPDILLQCLNLSIVSHATGCHVPGRI 180
 QY 181 TSNMVCAGGVPQDACCQDSGGPLVCGGLVQGLVSGSVGCGQDGPVYVYCKYVDW 240
 DB 181 TSNMVCAGGVPQDACCQDSGGPLVCGGLVQGLVSGSVGCGQDGPVYVYCKYVDW 240
 QY 241 IRMIMERN 248
 DB 241 IRMIMERN 248
 RESULT 9
 ABO33512

ID ABO33512 standard; protein; 248 AA.
XX
AC ABO33512;
XX
DT 17-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1303.
XX
KW Human; secreted and transmembrane protein; PRO; gene therapy; vaccine;
KW tissue typing; chromosome identification; vaccine.
OS Homo sapiens.
XX
PN US2003073129-A1.
XX
PD 17-APR-2003.
XX
XX 04-SEP-2001; 2001US-00946374.
XX
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 01-SEP-1998; 98US-0098749P.
PR 01-SEP-1998; 98US-0098750P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099536P.
PR 09-SEP-1998; 98US-0099596P.
PR 09-SEP-1998; 98US-0099598P.
PR 09-SEP-1998; 98US-0099602P.
PR 09-SEP-1998; 98US-0099642P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099792P.
PR 10-SEP-1998; 98US-0099808P.
PR 10-SEP-1998; 98US-0099812P.
PR 10-SEP-1998; 98US-0099815P.
PR 10-SEP-1998; 98US-0099818P.
PR 15-SEP-1998; 98US-0100385P.
PR 15-SEP-1998; 98US-0100388P.
PR 15-SEP-1998; 98US-0100390P.
PR 16-SEP-1998; 98US-0100584P.
PR 16-SEP-1998; 98US-0100627P.
PR 16-SEP-1998; 98US-0100661P.
PR 16-SEP-1998; 98US-0100662P.
PR 17-SEP-1998; 98US-0100664P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100710P.
PR 17-SEP-1998; 98US-0100711P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100848P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 18-SEP-1998; 98US-0101071P.
PR 22-SEP-1998; 98US-0101279P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101474P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101476P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101479P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102027P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106032P.
PR 28-OCT-1998; 98US-0106033P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106484P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 18-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-00218517.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 05-JAN-1999; 99WO-US000106.

```
PR 12-APR-1999; 99US-00284291.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 18-OCT-1999; 99US-00403297.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US000365.
PR 18-FEB-2000; 2000WO-US000434.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 01-JUN-2001; 2001US-00872035.
PR 14-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001US-00882636.
PR 20-JUN-2001; 2001WO-US018692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
XX Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
XX Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
XX Williams PM, Wood W;
XX
XX WPI: 2003-585292/55.
XX N-PSDB; ACD68024.
XX
XX Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the
XX preparation of a medicament for treating a condition responsive to PRO
XX polypeptide, and as therapeutic agents e.g. vaccines.
XX
XX Claim 12; Fig 108; 561pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (I), having at least 80% sequence identity to a sequence
XX
XX Query Match 100.0%; Score 1374; DB 7; Length 248;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-98;
XX Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MGLSIFLLLCVLGSLQATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWYLT 60
DB 1 MGLSIFLLLCVLGSLQATPKIFNGTECGRNSQPMQVGLFEGTSLRCGGVLIDHRWYLT 60
QY 61 AHCSGSRYWYRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPV 120
DB 61 AHCSGSRYWYRLGHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPV 120
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNPRNPFDPDLLOCLNLSIVSHATCHGVPGRI 180
DB 121 TSSVQPLPLPNDCATAGTECHVSGWGITNPRNPFDPDLLOCLNLSIVSHATCHGVPGRI 180
QY 181 TSNMVCAGGVPQDACQGDGGLVCGGV.LQGLVSWGSGVPCQGDGIPGVYTYICKYVDW 240
DB 181 TSNMVCAGGVPQDACQGDGGLVCGGV.LQGLVSWGSGVPCQGDGIPGVYTYICKYVDW 240
```

```
PR 23-SEP-1998; 98US-0101477P.
PR 23-SEP-1998; 98US-0101478P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101741P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101915P.
PR 24-SEP-1998; 98US-0101916P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102307P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102484P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0103314P.
PR 07-OCT-1998; 98US-0103315P.
PR 07-OCT-1998; 98US-0103328P.
PR 07-OCT-1998; 98US-0103395P.
PR 07-OCT-1998; 98US-0103396P.
PR 07-OCT-1998; 98US-0103401P.
PR 08-OCT-1998; 98US-0103633P.
PR 08-OCT-1998; 98US-0103678P.
PR 08-OCT-1998; 98US-0103679P.
PR 08-OCT-1998; 98US-0103711P.
PR 14-OCT-1998; 98US-0104257P.
PR 20-OCT-1998; 98US-0104987P.
PR 20-OCT-1998; 98US-0105000P.
PR 20-OCT-1998; 98US-0105002P.
PR 21-OCT-1998; 98US-0105104P.
PR 22-OCT-1998; 98US-0105169P.
PR 22-OCT-1998; 98US-0105266P.
PR 26-OCT-1998; 98US-0105693P.
PR 26-OCT-1998; 98US-0105694P.
PR 27-OCT-1998; 98US-0105807P.
PR 27-OCT-1998; 98US-0105881P.
PR 27-OCT-1998; 98US-0105882P.
PR 27-OCT-1998; 98US-0106062P.
PR 28-OCT-1998; 98US-0106023P.
PR 28-OCT-1998; 98US-0106029P.
PR 28-OCT-1998; 98US-0106030P.
PR 28-OCT-1998; 98US-0106033P.
PR 28-OCT-1998; 98US-0106178P.
PR 29-OCT-1998; 98US-0106248P.
PR 29-OCT-1998; 98US-0106384P.
PR 29-OCT-1998; 98US-0108500P.
PR 30-OCT-1998; 98US-0106464P.
PR 03-NOV-1998; 98US-0106856P.
PR 03-NOV-1998; 98US-0106902P.
PR 03-NOV-1998; 98US-0106905P.
PR 03-NOV-1998; 98US-0106919P.
PR 03-NOV-1998; 98US-0106932P.
PR 03-NOV-1998; 98US-0106934P.
PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108779P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.

PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 05-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006656.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 04-SEP-2001; 2001US-00946374.

PA (GETH ) GENENTECH INC.
XX
XX
PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;
XX WPI; 2003-555602/52.
DR N-PSDB; ADC18062.
XX
PT Novel isolated PRO polypeptides e.g. PRO1491 and PRO1571, useful in the
PT preparation of a medicament for treating a condition responsive to PRO
PT polypeptide, and as therapeutic agents e.g. vaccines.
XX
PS Claim 12; SEQ ID NO 194; 555pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The sequences are useful in the preparation of a
CC medicament for treating a condition responsive to a PRO polypeptide. The
CC polypeptides are useful in a number of functional biological assays, as
CC molecular weight markers for protein electrophoresis and as therapeutic

Query Match 100.0%; Score 1374; DB 7; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLSIFLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTA 60
Db 1 MGLSIFLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTA 60
Qy 61 AHCSGRYVVRNLGHSLSQLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVVR 120
```


Db 61 AHCSGSYWRLESHLSQDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLRPV 120
Qy 121 TSSVQPLPLPNDACATAGTECHVSGWGITNHPNFPDQLQCLNLSIVSHATCHGVYPGRI 180
Db 121 TSSVQPLPLPNDACATAGTECHVSGWGITNHPNFPDQLQCLNLSIVSHATCHGVYPGRI 180
Qy 181 TSNMVCAGGVPQDACQDGGPLVCGGVLOGLVSWGSGVPCQDGIPIGVYTYICKYVDW 240
Db 181 TSNMVCAGGVPQDACQDGGPLVCGGVLOGLVSWGSGVPCQDGIPIGVYTYICKYVDW 240
Qy 241 IRMIMRN 248
Db 241 IRMIMRN 248

RESULT 11
ADD70709
ID ADD70709 standard; protein; 248 AA.
XX AC ADD70709;
XX XX
Dt 15-JAN-2004 (first entry)
XX XX
De Human secreted/transmembrane protein PRO1303.
XX Human; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonlein-Henoch purpura; coeliac disease;
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX XX
OS Homo sapiens.
XX XX
PN US2003099625-A1.
XX XX
PD 29-MAY-2003.
XX XX
Pf 12-DEC-2001; 2001US-00015386.
XX XX
Pr 01-SEP-1998; 98US-0098716P.
Pr 01-SEP-1998; 98US-0098723P.
Pr 01-SEP-1998; 98US-0098749P.
Pr 01-SEP-1998; 98US-0098750P.
Pr 02-SEP-1998; 98US-0098803P.
Pr 02-SEP-1998; 98US-0098821P.
Pr 02-SEP-1998; 98US-0098843P.
Pr 09-SEP-1998; 98US-0099536P.
Pr 09-SEP-1998; 98US-0099596P.
Pr 09-SEP-1998; 98US-0099602P.
Pr 09-SEP-1998; 98US-0099642P.
Pr 10-SEP-1998; 98US-0099741P.
Pr 10-SEP-1998; 98US-0099754P.
Pr 10-SEP-1998; 98US-0099763P.
Pr 10-SEP-1998; 98US-0099792P.
Pr 10-SEP-1998; 98US-0099808P.
Pr 10-SEP-1998; 98US-0099812P.
Pr 10-SEP-1998; 98US-0099815P.
Pr 15-SEP-1998; 98US-0100385P.
Pr 15-SEP-1998; 98US-0100388P.
Pr 15-SEP-1998; 98US-0100390P.
Pr 16-SEP-1998; 98US-0100584P.
Pr 16-SEP-1998; 98US-0100627P.
Pr 16-SEP-1998; 98US-0100661P.
Pr 16-SEP-1998; 98US-0100662P.
Pr 16-SEP-1998; 98US-0100664P.
Pr 17-SEP-1998; 98US-0100683P.
Pr 17-SEP-1998; 98US-0100684P.
Pr 17-SEP-1998; 98US-0100710P.
Pr 17-SEP-1998; 98US-0100711P.

Pr 17-SEP-1998; 98US-0100919P.
Pr 17-SEP-1998; 98US-0100930P.
Pr 18-SEP-1998; 98US-0100848P.
Pr 18-SEP-1998; 98US-0100849P.
Pr 18-SEP-1998; 98US-0101014P.
Pr 18-SEP-1998; 98US-0101068P.
Pr 18-SEP-1998; 98US-0101071P.
Pr 22-SEP-1998; 98US-0101279P.
Pr 23-SEP-1998; 98US-0101471P.
Pr 23-SEP-1998; 98US-0101472P.
Pr 23-SEP-1998; 98US-0101474P.
Pr 23-SEP-1998; 98US-0101475P.
Pr 23-SEP-1998; 98US-0101476P.
Pr 23-SEP-1998; 98US-0101477P.
Pr 24-SEP-1998; 98US-0101479P.
Pr 24-SEP-1998; 98US-0101738P.
Pr 24-SEP-1998; 98US-0101741P.
Pr 24-SEP-1998; 98US-0101743P.
Pr 24-SEP-1998; 98US-0101915P.
Pr 24-SEP-1998; 98US-0101916P.
Pr 29-SEP-1998; 98US-0102207P.
Pr 29-SEP-1998; 98US-0102240P.
Pr 29-SEP-1998; 98US-0102307P.
Pr 29-SEP-1998; 98US-0102330P.
Pr 29-SEP-1998; 98US-0102331P.
Pr 30-SEP-1998; 98US-0102484P.
Pr 30-SEP-1998; 98US-0102487P.
Pr 30-SEP-1998; 98US-0102570P.
Pr 01-OCT-1998; 98US-0102571P.
Pr 01-OCT-1998; 98US-0102684P.
Pr 01-OCT-1998; 98US-0102687P.
Pr 02-OCT-1998; 98US-0102965P.
Pr 06-OCT-1998; 98US-0103258P.
Pr 06-OCT-1998; 98US-0103499P.
Pr 07-OCT-1998; 98US-0103314P.
Pr 07-OCT-1998; 98US-0103315P.
Pr 07-OCT-1998; 98US-0103328P.
Pr 07-OCT-1998; 98US-0103395P.
Pr 07-OCT-1998; 98US-0103396P.
Pr 07-OCT-1998; 98US-0103401P.
Pr 08-OCT-1998; 98US-0103633P.
Pr 08-OCT-1998; 98US-0103678P.
Pr 08-OCT-1998; 98US-0103679P.
Pr 08-OCT-1998; 98US-0103711P.
Pr 14-OCT-1998; 98US-0104257P.
Pr 20-OCT-1998; 98US-0104987P.
Pr 20-OCT-1998; 98US-0105000P.
Pr 20-OCT-1998; 98US-0105002P.
Pr 22-OCT-1998; 98US-0105169P.
Pr 22-OCT-1998; 98US-0105266P.
Pr 26-OCT-1998; 98US-0105693P.
Pr 26-OCT-1998; 98US-0105694P.
Pr 27-OCT-1998; 98US-0105807P.
Pr 27-OCT-1998; 98US-0105881P.
Pr 27-OCT-1998; 98US-0105882P.
Pr 27-OCT-1998; 98US-0106062P.
Pr 28-OCT-1998; 98US-0106023P.
Pr 28-OCT-1998; 98US-0106029P.
Pr 28-OCT-1998; 98US-0106030P.
Pr 28-OCT-1998; 98US-0106032P.
Pr 28-OCT-1998; 98US-0106033P.
Pr 28-OCT-1998; 98US-0106178P.
Pr 29-OCT-1998; 98US-0106248P.
Pr 29-OCT-1998; 98US-0106384P.
Pr 29-OCT-1998; 98US-0108500P.
Pr 30-OCT-1998; 98US-0106464P.
Pr 03-NOV-1998; 98US-0106856P.
Pr 03-NOV-1998; 98US-0106902P.
Pr 03-NOV-1998; 98US-0106905P.
Pr 03-NOV-1998; 98US-0106919P.
Pr 03-NOV-1998; 98US-0106932P.
Pr 03-NOV-1998; 98US-0106934P.

PR 10-NOV-1998; 98US-0107783P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108775P.
PR 17-NOV-1998; 98US-0108787P.
PR 17-NOV-1998; 98US-0108788P.
PR 17-NOV-1998; 98US-0108801P.
PR 17-NOV-1998; 98US-0108802P.
PR 17-NOV-1998; 98US-0108806P.
PR 17-NOV-1998; 98US-0108807P.
PR 17-NOV-1998; 98US-0108867P.
PR 17-NOV-1998; 98US-0108925P.
PR 18-NOV-1998; 98US-0108848P.
PR 18-NOV-1998; 98US-0108849P.
PR 18-NOV-1998; 98US-0108850P.
PR 18-NOV-1998; 98US-0108851P.
PR 18-NOV-1998; 98US-0108852P.
PR 18-NOV-1998; 98US-0108858P.
PR 18-NOV-1998; 98US-0108904P.
PR 22-DEC-1998; 98US-0113296P.
PR 30-DEC-1998; 98US-0114223P.
PR 05-JAN-1999; 99WO-US000106.
PR 16-APR-1999; 99US-0129674P.
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021194.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004342.
PR 24-FEB-2000; 2000WO-US005841.
PR 02-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US006884.
PR 24-AUG-2000; 2000WO-US023328.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021086.
PR 09-JUL-2001; 2001WO-US021735.
PR 04-SEP-2001; 2001US-00946374.
XX (GETH) GENENTECH INC.
PA Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,
XX Gao W, Goddard A, Godowski J, Grimaldi JC, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;
XX WPI: 2003-874602/81.
DR N-PSDB; ADD70708.
XX
PT Novel isolated PRO polypeptides e.g., PRO1130, PRO1275, PRO1418, PRO1555,
PT PRO1787 affect glucose or free fatty acid (FFA) uptake by skeletal muscle
PT cells and are useful for treating diabetes or hyper- or hypo-insulinemia.
XX
PS Claim 12; SEQ ID NO 194; 553pp; English.
XX
CC The invention relates to an isolated PRO polypeptide (secreted or

Query Match 100.0%; Score 1374; DB 7; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98; Indels 0; Gaps 0;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLSIFLLCVLGSLQAATPKIFNGTECGRNSQFQVGLFEGTSLRCGGVLIDHRWLTA 60
Db 1 MGLSIFLLCVLGSLQAATPKIFNGTECGRNSQFQVGLFEGTSLRCGGVLIDHRWLTA 60
Qy 61 AHCSGSRYWVRLGEHSLQSLDWTETQIRHSGFSVTHPGYLGASTSHEHDLRLRLPRV 120
Db 61 AHCSGSRYWVRLGEHSLQSLDWTETQIRHSGFSVTHPGYLGASTSHEHDLRLRLPRV 120
Qy 121 TSSVQPLPLNDCATAGTECHVSGWGITNHPRPFPDLLQCLNLSIVSHATCHGVYPGRI 180
Db 121 TSSVQPLPLNDCATAGTECHVSGWGITNHPRPFPDLLQCLNLSIVSHATCHGVYPGRI 180
Qy 181 TSNMVCAGGVPGQDACQDSGGPLVCGVQLQGLVSGSVGPGQDGPVYTYICKYVDW 240
Db 181 TSNMVCAGGVPGQDACQDSGGPLVCGVQLQGLVSGSVGPGQDGPVYTYICKYVDW 240
Qy 241 IRMIMRNN 248
Db 241 IRMIMRNN 248
RESULT 12
ADD33786
ID ADD33786 standard; protein; 248 AA.
XX AC ADD33786;
XX DT 15-JAN-2004 (first entry)
XX DE Human secreted/transmembrane protein PRO1303.
XX KW Human; secreted protein; transmembrane protein; PRO; tumour;
KW immune response; cardiac insufficiency disorder; calcium flux;
KW umbilical vein endothelial cell; bone disorder; cartilage disorder;
KW arthritis; wound healing; diabetes; skeletal muscle cells; obesity;
KW Berger disease; nephropathy; Schonelein-Henoch purpura; coeliac disease;
KW dermatitis; herpeticiformis; Crohn's disease; thalassaemia.
XX OS Homo sapiens.
XX PN US2003083462-A1.
XX PD 01-MAY-2003.
XX PF 10-DEC-2001; 2001US-00013913.
XX PR 05-JAN-1999; 99WO-US000106.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 15-SEP-1999; 99WO-US021194.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 02-DEC-1999; 99WO-US028551.
XX PR 16-DEC-1999; 99WO-US030095.
XX PR 05-JAN-2000; 2000WO-US000219.
XX PR 06-JAN-2000; 2000WO-US000376.
XX PR 11-FEB-2000; 2000WO-US003565.
XX PR 18-FEB-2000; 2000WO-US004342.
XX PR 24-FEB-2000; 2000WO-US005004.
XX PR 02-MAR-2000; 2000WO-US005841.
XX PR 15-MAR-2000; 2000WO-US006884.
XX PR 17-MAY-2000; 2000WO-US013705.
XX PR 22-MAY-2000; 2000WO-US014042.
XX PR 30-MAY-2000; 2000WO-US014941.
XX PR 23-AUG-2000; 2000WO-US015264.
XX PR 24-AUG-2000; 2000WO-US023328.
XX PR 08-NOV-2000; 2000WO-US030952.
XX PR 10-NOV-2000; 2000WO-US030873.
XX PR 01-DEC-2000; 2000WO-US032678.
XX PR 28-FEB-2001; 2001WO-US006520.

PR 23-SEP-1998; 98US-0101471P.
 PR 23-SEP-1998; 98US-0101472P.
 PR 23-SEP-1998; 98US-0101473P.
 PR 23-SEP-1998; 98US-0101474P.
 PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101476P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 23-SEP-1998; 98US-0101478P.
 PR 23-SEP-1998; 98US-0101479P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101741P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101915P.
 PR 24-SEP-1998; 98US-0101916P.
 PR 24-SEP-1998; 98US-0102207P.
 PR 24-SEP-1998; 98US-0102240P.
 PR 29-SEP-1998; 98US-0102307P.
 PR 29-SEP-1998; 98US-0102330P.
 PR 29-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102484P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.
 PR 02-OCT-1998; 98US-0102965P.
 PR 06-OCT-1998; 98US-0103258P.
 PR 06-OCT-1998; 98US-0103449P.
 PR 07-OCT-1998; 98US-0103314P.
 PR 07-OCT-1998; 98US-0103315P.
 PR 07-OCT-1998; 98US-0103328P.
 PR 07-OCT-1998; 98US-0103395P.
 PR 07-OCT-1998; 98US-0103398P.
 PR 07-OCT-1998; 98US-0103401P.
 PR 08-OCT-1998; 98US-0103633P.
 PR 08-OCT-1998; 98US-0103678P.
 PR 08-OCT-1998; 98US-0103679P.
 PR 08-OCT-1998; 98US-0103711P.
 PR 14-OCT-1998; 98US-0104257P.
 PR 20-OCT-1998; 98US-0104987P.
 PR 20-OCT-1998; 98US-0105000P.
 PR 20-OCT-1998; 98US-0105002P.
 PR 21-OCT-1998; 98US-0105104P.
 PR 22-OCT-1998; 98US-0105169P.
 PR 22-OCT-1998; 98US-0105266P.
 PR 26-OCT-1998; 98US-0105693P.
 PR 26-OCT-1998; 98US-0105694P.
 PR 27-OCT-1998; 98US-0105807P.
 PR 27-OCT-1998; 98US-0105881P.
 PR 27-OCT-1998; 98US-0105882P.
 PR 27-OCT-1998; 98US-0106062P.
 PR 28-OCT-1998; 98US-0106023P.
 PR 28-OCT-1998; 98US-0106029P.
 PR 28-OCT-1998; 98US-0106030P.
 PR 28-OCT-1998; 98US-0106032P.
 PR 28-OCT-1998; 98US-0106033P.
 PR 28-OCT-1998; 98US-0106178P.
 PR 28-OCT-1998; 98US-0106248P.
 PR 29-OCT-1998; 98US-0106384P.
 PR 29-OCT-1998; 98US-0108500P.
 PR 30-OCT-1998; 98US-0106464P.
 PR 03-NOV-1998; 98US-0106856P.
 PR 03-NOV-1998; 98US-0106902P.
 PR 03-NOV-1998; 98US-0106905P.
 PR 03-NOV-1998; 98US-0106919P.
 PR 03-NOV-1998; 98US-0106932P.
 PR 10-NOV-1998; 98US-0106934P.
 PR 17-NOV-1998; 98US-0107783P.
 PR 17-NOV-1998; 98US-0108775P.
 PR 17-NOV-1998; 98US-0108779P.
 PR 17-NOV-1998; 98US-0108788P.
 PR 17-NOV-1998; 98US-0108801P.
 PR 17-NOV-1998; 98US-0108802P.
 PR 17-NOV-1998; 98US-0108806P.

PR 17-NOV-1998; 98US-0108907P.
 PR 17-NOV-1998; 98US-0108867P.
 PR 17-NOV-1998; 98US-0108925P.
 PR 18-NOV-1998; 98US-0108848P.
 PR 18-NOV-1998; 98US-0108849P.
 PR 18-NOV-1998; 98US-0108850P.
 PR 18-NOV-1998; 98US-0108851P.
 PR 18-NOV-1998; 98US-0108852P.
 PR 18-NOV-1998; 98US-0108858P.
 PR 18-NOV-1998; 98US-0108904P.
 PR 22-DEC-1998; 98US-0113296P.
 PR 30-DEC-1998; 98US-0114223P.
 PR 05-JAN-1999; 99WO-US000106.
 PR 16-APR-1999; 99US-0129674P.
 PR 23-JUN-1999; 99US-0141037P.
 PR 20-JUL-1999; 99US-0144758P.
 PR 26-JUL-1999; 99US-0145698P.
 PR 01-SEP-1999; 99WO-US020111.
 PR 15-SEP-1999; 99WO-US021194.
 PR 29-OCT-1999; 99US-0152506P.
 PR 30-NOV-1999; 99WO-US028313.
 PR 02-DEC-1999; 99WO-US028551.
 PR 16-DEC-1999; 99WO-US030095.
 PR 06-JAN-2000; 2000WO-US000219.
 PR 11-FEB-2000; 2000WO-US000376.
 PR 18-FEB-2000; 2000WO-US003565.
 PR 24-FEB-2000; 2000WO-US004342.
 PR 02-MAR-2000; 2000WO-US005004.
 PR 15-MAR-2000; 2000WO-US005841.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-MAR-2001; 2001WO-US006666.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 04-SEP-2001; 2001US-00946374.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
 PI Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
 PI Williams PM, Wood WI;
 XX
 DR WPI; 2003-708344/67.
 DR N-PSDB; ADD70231.
 XX
 PT Novel isolated PRO polypeptide useful for tissue typing, modulating
 PT biological activity of cell, as molecular weight markers in protein
 PT electrophoresis, for treating arthritis, tumor.
 XX
 PS Claim 12; SEQ ID NO 194; 549pp; English.
 XX
 CC The invention relates to an isolated PRO polypeptide (secreted or

Query Match 100.0%; Score 1374; DB 7; Length 248;
 Best Local Similarity 100.0%; Pred. No. 3.8e-98;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLLCVLGLSQANTPKIFNCTGCRNSQPMQVGLFEGTSIRCGGLIDHFWLTA 60
 |||||
 DB 1 MGLSIFLLLCVLGLSQANTPKIFNCTGCRNSQPMQVGLFEGTSIRCGGLIDHFWLTA 60
 |||||

PR 15-SEP-1998; 98US-0100390P.
 PR 16-SEP-1998; 98US-0100584P.
 PR 16-SEP-1998; 98US-0100627P.
 PR 16-SEP-1998; 98US-0100661P.
 PR 16-SEP-1998; 98US-0100662P.
 PR 16-SEP-1998; 98US-0100664P.
 PR 17-SEP-1998; 98US-0100683P.
 PR 17-SEP-1998; 98US-0100684P.
 PR 17-SEP-1998; 98US-0100710P.
 PR 17-SEP-1998; 98US-0100711P.
 PR 17-SEP-1998; 98US-0100919P.
 PR 17-SEP-1998; 98US-0100930P.
 PR 18-SEP-1998; 98US-0100848P.
 PR 18-SEP-1998; 98US-0100849P.
 PR 18-SEP-1998; 98US-0101014P.
 PR 18-SEP-1998; 98US-0101068P.
 PR 18-SEP-1998; 98US-0101071P.
 PR 22-SEP-1998; 98US-0101279P.
 PR 23-SEP-1998; 98US-0101471P.
 PR 23-SEP-1998; 98US-0101472P.
 PR 23-SEP-1998; 98US-0101474P.
 PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101476P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 23-SEP-1998; 98US-0101479P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101741P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101915P.
 PR 24-SEP-1998; 98US-0101916P.
 PR 24-SEP-1998; 98US-0102207P.
 PR 23-SEP-1998; 98US-0102240P.
 PR 23-SEP-1998; 98US-0102307P.
 PR 23-SEP-1998; 98US-0102330P.
 PR 23-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102484P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.
 PR 02-OCT-1998; 98US-0102965P.
 PR 06-OCT-1998; 98US-0103258P.
 PR 06-OCT-1998; 98US-0103449P.
 PR 07-OCT-1998; 98US-0103314P.
 PR 07-OCT-1998; 98US-0103315P.
 PR 07-OCT-1998; 98US-0103328P.
 PR 07-OCT-1998; 98US-0103395P.
 PR 07-OCT-1998; 98US-0103396P.
 PR 08-OCT-1998; 98US-0103401P.
 PR 08-OCT-1998; 98US-0103633P.
 PR 08-OCT-1998; 98US-0103678P.
 PR 08-OCT-1998; 98US-0103799P.
 PR 08-OCT-1998; 98US-0103711P.
 PR 14-OCT-1998; 98US-0104257P.
 PR 20-OCT-1998; 98US-0104987P.
 PR 20-OCT-1998; 98US-0105000P.
 PR 20-OCT-1998; 98US-0105002P.
 PR 21-OCT-1998; 98US-0105104P.
 PR 22-OCT-1998; 98US-0105169P.
 PR 22-OCT-1998; 98US-0105266P.
 PR 26-OCT-1998; 98US-0105693P.
 PR 26-OCT-1998; 98US-0105694P.
 PR 27-OCT-1998; 98US-0105807P.
 PR 27-OCT-1998; 98US-0105881P.
 PR 27-OCT-1998; 98US-0105882P.
 PR 27-OCT-1998; 98US-0106062P.
 PR 28-OCT-1998; 98US-0106023P.
 PR 28-OCT-1998; 98US-0106029P.
 PR 28-OCT-1998; 98US-0106030P.
 PR 28-OCT-1998; 98US-0106032P.
 PR 28-OCT-1998; 98US-0106033P.
 PR 28-OCT-1998; 98US-0106178P.
 PR 28-OCT-1998; 98US-0106179P.

PR 29-OCT-1998; 98US-0106248P.
 PR 29-OCT-1998; 98US-0106384P.
 PR 29-OCT-1998; 98US-0108500P.
 PR 30-OCT-1998; 98US-0106464P.
 PR 03-NOV-1998; 98US-0106856P.
 PR 03-NOV-1998; 98US-0108902P.
 PR 03-NOV-1998; 98US-0108905P.
 PR 03-NOV-1998; 98US-0106919P.
 PR 03-NOV-1998; 98US-0106932P.
 PR 03-NOV-1998; 98US-0106934P.
 PR 10-NOV-1998; 98US-0107783P.
 PR 17-NOV-1998; 98US-0108775P.
 PR 17-NOV-1998; 98US-0108779P.
 PR 17-NOV-1998; 98US-0108787P.
 PR 17-NOV-1998; 98US-0108788P.
 PR 17-NOV-1998; 98US-0108801P.
 PR 17-NOV-1998; 98US-0108802P.
 PR 17-NOV-1998; 98US-0108806P.
 PR 17-NOV-1998; 98US-0108807P.
 PR 17-NOV-1998; 98US-0108867P.
 PR 17-NOV-1998; 98US-0108925P.
 PR 18-NOV-1998; 98US-0108848P.
 PR 18-NOV-1998; 98US-0108849P.
 PR 18-NOV-1998; 98US-0108850P.
 PR 18-NOV-1998; 98US-0108851P.
 PR 18-NOV-1998; 98US-0108852P.
 PR 18-NOV-1998; 98US-0108858P.
 PR 22-DEC-1998; 98US-0108904P.
 PR 22-DEC-1998; 98US-0113296P.
 PR 05-JAN-1999; 99WO-US000106.
 PR 16-APR-1999; 99US-0129674P.
 PR 23-JUN-1999; 99US-0141037P.
 PR 26-JUL-1999; 99US-014758P.
 PR 01-SEP-1999; 99WO-US020111.
 PR 15-SEP-1999; 99WO-US021194.
 PR 23-OCT-1999; 99US-0162506P.
 PR 30-NOV-1999; 99WO-US028313.
 PR 02-DEC-1999; 99WO-US028551.
 PR 16-DEC-1999; 99WO-US030095.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 28-FEB-2001; 2001WO-US000520.
 PR 01-MAR-2001; 2001WO-US006666.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 04-SEP-2001; 2001US-00946374.

(GETH) GENENTECH INC.

XX Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;
 XX Gao W, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PI Pan J, Faoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
 PI Williams PM, Wood WI;
 XX WPI; 2003-786999/74.
 DR

DR N-PSDB; ADD39308.

XX Novel isolated PRO polypeptide useful for tissue typing, modulating
PT biological activity of cell, as molecular weight markers in protein
PT electrophoresis, for treating arthritis, tumor.

PS Claim 12; SEQ ID NO 194; 550pp; English.

XX The invention relates to an isolated PRO polypeptide (secreted or
CC

Query Match 100.0%; Score 1374; DB 7; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.8e-98;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLT 60

Db 1 MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLT 60

QY 61 AHCSGSRYWVRLGEHSLQOLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120

Db 61 AHCSGSRYWVRLGEHSLQOLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLRLPVRV 120

QY 121 TSSVQPLPLPNDCATAGTECHVSGWGI TNHPRNPPDLLQCLNLSIVSHATCHGVYPGRI 180

Db 121 TSSVQPLPLPNDCATAGTECHVSGWGI TNHPRNPPDLLQCLNLSIVSHATCHGVYPGRI 180

QY 181 TSNMVCAGGVPGQDACQDSDGGLVCGGVLOGLVSWGSGVPCGQDGI PGVYTYICKYVDW 240

Db 181 TSNMVCAGGVPGQDACQDSDGGLVCGGVLOGLVSWGSGVPCGQDGI PGVYTYICKYVDW 240

QY 241 IRMIMRNN 248

Db 241 IRMIMRNN 248

Search completed: June 16, 2004, 20:32:43

Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:30:31 ; Search time 20 seconds
(without alignments)
1192.775 Million cell updates/sec

Title: US-10-015-385A-194

Perfect score: 1374
Sequence: 1 MGSLIFLLCVLGLSQAATP.....GVYYICKYVDIMRMNRN 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR 78:*
1: Pirl1.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	622.5	45.3	260	2	neuropsin - mouse
2	569.5	41.4	248	2	trypsin EC 3.4.21
3	567	41.3	248	2	trypsin EC 3.4.21
4	561	40.8	231	1	trypsin EC 3.4.21
5	555	40.4	243	2	trypsin EC 3.4.21
6	555	40.4	253	2	serine proteinase
7	553	40.2	246	2	trypsin EC 3.4.21
8	552	40.2	229	1	trypsin EC 3.4.21
9	550	40.0	247	2	trypsin EC 3.4.21
10	546	39.7	247	1	trypsin EC 3.4.21
11	545	39.7	246	1	trypsin EC 3.4.21
12	544.5	39.6	247	2	trypsin EC 3.4.21
13	543	39.5	246	1	trypsin EC 3.4.21
14	540	39.3	238	2	trypsin EC 3.4.21
15	533.5	38.8	247	2	trypsin EC 3.4.21
16	532	38.7	246	1	trypsin EC 3.4.21
17	527.5	38.4	261	2	tissue kallikrein
18	524	38.1	232	1	tissue kallikrein
19	520.5	37.9	261	2	tissue kallikrein
20	519	37.8	246	2	tissue kallikrein
21	518	37.7	246	2	tissue kallikrein
22	518	37.7	261	2	tissue kallikrein
23	513.5	37.4	231	2	trypsin EC 3.4.21
24	513.5	37.4	239	2	trypsin EC 3.4.21
25	513.5	37.2	261	2	tissue kallikrein
26	510.5	37.0	261	1	7S nerve growth fa
27	508	37.0	242	2	trypsin EC 3.4.21
28	508	37.0	242	2	trypsin EC 3.4.21
29	503	36.6	260	2	tissue kallikrein

30 503 36.6 263 2 S15686 tissue kallikrein
31 502 36.5 259 2 B31136 tissue kallikrein
32 501 36.5 247 1 B25852 trypsin EC 3.4.21
33 499.5 36.4 241 2 S39048 trypsin EC 3.4.21
34 496.5 36.1 261 1 TRMSMS tissue kallikrein
35 496 36.1 242 2 S49489 trypsin EC 3.4.21
36 495.5 36.1 257 2 S33772 tissue kallikrein
37 495.5 36.1 261 2 S01971 tissue kallikrein
38 494.5 36.0 247 1 A25852 tissue kallikrein
39 493 35.9 262 1 KQHU tissue kallikrein
40 491.5 35.8 240 2 S39047 trypsin EC 3.4.21
41 491 35.7 256 1 NGMSA 7S nerve growth fa
42 490.5 35.7 229 1 TRDFS trypsin EC 3.4.21
43 490.5 35.7 261 2 A41020 tissue kallikrein
44 489 35.6 244 2 A44284 tissue kallikrein
45 485.5 35.3 261 2 JE0236 tissue kallikrein

ALIGNMENTS

RESULT 1

I56559
neuropsin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 21-Jul-2000
C:Accession: I56559
R:Chen, Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Ni J. Neurosci. 15, 5088-5097, 1995
A:Title: Expression and activity-dependent changes of a novel limbic-serine protease
A:Reference number: I56559; MUID:95348817; PMID:7623137
A:Accession: I56559
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-260 <RES>
A:Cross-references: GB:D30785; NID:G1648847; PIDN:BAA06451.1; PID:G1020091
C:Superfamily: trypsin; trypsin homology
F:33-252/Domain: trypsin homology <TRY>

Query Match 45.3%; Score 622.5; DB 2; Length 260;
Best Local Similarity 49.0%; Pred. No. 1.4e-43;
Matches 121; Conservative 28; Mismatches 93; Indels 5; Gaps 3;

Qy 5 IFLLLCV---LGLSQATPKIFNGTEGSRNSQPNQVGLFEGTSIRCCGVLIDHWLTAA 61
Db 13 ILLLLFMGAWAGLTRAQGSKILEGECIPHSQPNQAAALFQGERLICGGVLGDRWLTAA 72
Qy 62 HCSGRYVVRIGHSLSQLDWTEQIRHSGFVTHPGVLGAS-TSHEHDLRLRLRPVRV 120
Db 73 HCKKQKYSVRLGDSLSQSDQPEIQVAQSIQHPCTYNSNPEDSHDMLIRLQNSANL 132
Qy 121 TSSVQPLPLPNDCACTAGTECHVSGWGTNHPNPFDPDLQCLNLSIVSHATCHGVYFGR 180
Db 133 GDKVKPQLANLCPKVGKCIISGWGTWTSPOENFNTLNCAEVKIYSQNKCEAYFGKI 192
Qy 181 TSNMVCAGVPGQDACCGDSGGLVCGVLOGLVSWSGVPGCGQDGPVYTYICKYVDW 240
Db 193 TEGMVCAGSSNGADTCQDSGGPLVCDMLQGITWSGS-DPCGRKPEKPGVYTKICRYTWT 251
Qy 241 IRIMLRN 247
Db 252 IKKTMND 258

RESULT 2

S55066
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
N:Alternate names: trypsinogen II
C:Species: Gallus gallus (chicken)
C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
C:Accession: S55066; S72347
R:Wang, K.; Gan, L.; Lee, I.; Hood, L.
Biochem. J. 307, 471-479, 1995

A;Title: Isolation and characterization of the chicken trypsinogen gene family.
A;Reference number: S55065; MUID:95251611; PMID:7733885
A;Accession: S55066
A;Molecule type: mRNA
A;Residues: 1-248 <WAN1>
A;Cross-references: EMBL:U15157; NID:G603906; PIDN:AAA79914.1; PID:G603907
A;Experimental source: clone 2-P29
A;Accession: S72347
A;Molecule type: DNA
A;Residues: 1-248 <WAN2>
A;Cross-references: EMBL:U15157; NID:G603906; PIDN:AAA79914.1; PID:G603907
A;Experimental source: clone 2-P29
C;Superfamily: trypsin; trypsin homology
E;1-16/Domain: signal sequence #status predicted <SIG>
E;17-28/Domain: activation peptide #status predicted <APT>
F;26-248/Product: trypsin II #status predicted <WAR>
F;26-241/Domain: trypsin homology <TRY>
F;65,109,202/Active site:His, Asp, Ser #status predicted

Query Match.	41.4%;	Score	569.5;	DB 2;	Length	248;			
Best local Similarity	47.5%;	Fred. NO.	2.9e-39;						
Matches	116;	Conservative	39;	Mismatches	78;	Indels	11;	Gaps	6;
Qy	5	I F I L L C V I G L S - - - - Q A A T P K I F N G T E C G R N S O P W Q V G L F E G T S L R C G V L I D H R W V I T	59						
		: : : : : : : : : : : : : : : : : : : :							
Dd	4	L F L I L S C L G A A V A P F G G A D D K I V G G Y T C P E H S V P V Q V S L N S G Y H F - C G G S L I N S Q W V L S	62						
		: : : : : : : : : : : : : : : : : : : :							
Qy	60	A A H C S G Y W Y R I G E H S L S Q L D W T E Q I R H S G F S V T H P C Y L G A S T S H E D L L R L L R L P V R	119						
		: : : : : : : : : : : : : : :							
Dd	63	A A H C Y K S I Q I V R L G E Y N I D V Q E D S E V R S S V I R H P K Y - - S S I T L N D I M L I K A S A V E	120						
		: : : : : : : : : : : : : : :							
Qy	120	V T S S V Q P L P E N D C A T A G E C H V S G W G T T N H R N P F P D L L Q C L N L S I V S H A T C H G V P E G R	179						
		: : : : : : : : : : : : : : :							
Dd	121	Y S A D I Q P I A L P S S C A K A G T E C L I S G W G T L S N G Y N Y P E L L Q C L N A P I L S D Q E C Q A Y P G D	180						
		: : : : : : : : : : : : : : :							
Qy	180	I T S N M V C A G G V P - G O D A C G G S G G P L V C G G V L Q G L A V S W G S V C P C Q D G I P G V Y T V I C K X V	238						
		: : : : : : : : : : : : : : :							
Dd	181	I T S N M I C V G F L E G G K D S C Q G S G G P V N G E L Q G I V S W G - I G - C A L K G P G V Y T T K V C N V T	238						
		: : : : : : : : : : : : : : :							
Qy	239	DWIR	242						
		:							
Dd	239	DWIO	242						

```

RESULT 3
S55067
trypsin (EC 3.4.21.4) I precursor, pancreatic - chicken
N:Alternate names: trypsinogen I
C:Species: Gallus gallus (chicken)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 21-Jul-2000
C;Accession: S55067; S72345; S55065; S72346; S71155
R;Wang, K.; Gan, L.; Lee, I.; Hood, L.
B;Biochem. J. 307, 471-479, 1995
A;Title: Isolation and Characterization of the chicken trypsinogen gene family
A;Reference number: S55065; MOID:95251611; PMID:7733885
A;Accession: S55067
A;Molecule type: mRNA
A;Residues: 1-248 <WAN1>
A;Cross-references: EMBL:U15156; NID:g603904
A;Experimental source: clone 1-p38
A;Accession: S72345
A;Molecule type: DNA
A;Residues: 1-248 <WAN2>
A;Experimental source: clone 1-p38
A;Accession: S55065
A;Molecule type: mRNA
A;Residues: 1-9,'V',11-12,'T',14-102,'A',104-214,'I',216-248 <WAN3>
A;Cross-references: EMBL:U15155; NID:g603902; PIDN:AAA79912.1; PID:g603903
A;Experimental source: clone P1
A;Accession: S72346
A;Molecule type: DNA
A;Residues: 1-9,'V',11-12,'T',14-102,'A',104-214,'I',216-248 <WAN4>

```

A;Cross-references: GB:U15155; NID:g603902; PIDN:AA79912.1; PID:g603903
A;Experimental source: Clone P1
R;Wang, X.
Submitted to the EMBL Data Library, September 1994
A;Reference number: S71155
A;Accession: S71155
A;Molecule type: mRNA
A;Residues: 1-102, 'A', 104-248 <WAS>
A;Cross-references: EMBL:U15156; NID:g603904; PIDN:AA79913.1; PID:g603905
A;Experimental source: clone 1-P38
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
E;1-15/Domain: signal sequence #status predicted <SIG>
F;16-25/Domain: activation peptide #status predicted <APT>
F;26-248/Product: trypsin I #status predicted <WAT>
F;26-241/Domain: trypsin homology <TRY>
F;65,109,202/Active site: His, Asp, Ser #status predicted

Query Match	41.3%	Score	567	DB 2	Length	248			
Best Local Similarity	45.5%	Pred. No.	4.6e-39						
Matches	115	Conservative	38	Mismatches	88	Indels	12	Gaps	6
Qy	3	LSIFLLCVLG-----LSQAATPKIFNGTECGRNSQPQVGLFEGTSLRCGGVLIDHRW	56						
Db	1	MKFLVVAFLGAVAVAFPSIDEDDDKIVGGYSCARSAPYQVSLNSGYHF--CGGSLISSQW	59						
Qy	57	VLTAHCGSGYVWVRLGEHSLSQLDWTQEI RHSGFVTHPGVLGASTSHEHDLRLRLRL	116						
Db	60	VLSAAHCYKSIQVKGLEYINLAADGSGQTTSISSKVI RHSGY--NSNTLNNDIMLIKUSK	117						
Qy	117	PVRTSSVQPLPLPNDCATAGTECHVSGWGINTHPRNPFPLDLQCLNLSIVSHATCHGVY	176						
Db	118	AATLSYVNTVPELPTSCVTAGTCTCLISGWNGLTSGSGLYPDVLQCLNAPVLSSQCSAY	177						
Qy	177	PGRITTSNMVCAGGV--PGODACQSGGGLVCGGVLOGLVSGWSGVPCCGQDGIIPGVVYTC	235						
Db	178	PGRITTSNMICLYLNGGKDKSCQSGGPGPVNCGLQGFVSWG--IG-CAKQGPVGYTKVC	235						
Qy	236	KYVDWIRMIMRNN	248						
Db	236	NYVSWIKTTMSGN	248						

```

RESULT 4
TRPGTR
trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
N:Contains: trypsinogen
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 24-Apr-1984 #sequence revision 24-Apr-1984 #text_change 31-Mar-2000
C/Accession: A90641; A90368; A00947
R/Charles, M.; Roveiy, M.; Guidoni, A.; Desnuelle, P.
Biochim. Biophys. Acta 69, 115-129, 1963
A>Title: Su le trypsinogene et la trypsine de porc.
A/Reference number: A90641
A/Accession: A90641
A/Molecule type: protein
A/Residues: 1-10 <CHA>
R/Hermidson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 12, 3146-3153, 1973
A>Title: Determination of the amino acid sequence of porcine trypsin by sequenator a
A/Reference number: A90368; MUID:73258692; PMID:4738933
A/Accession: A90368
A/Molecule type: protein
A/Residues: 9-231 <HER>
A/Note: at position 20, Ile and Val occur alternatively
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase
F.1-231/Product: trypsinogen #status experimental <ZYM>
F.1-8/Domain: activation peptide #status experimental <APT>
F.9-231/Product: trypsin #status experimental <MAT>
F.9-224/Domain: trypsin homology <TRY>
F.115-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted
F.48-92,185/Active site: His, Asp, Ser #status predicted

```

```
F;60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Query Match 40.4%; Score 561; DB 1; Length 231;
Best Local Similarity 46.7%; Pred. No. 1.3e-38;
Matches 107; Conservative 39; Mismatches 77; Indels 6; Gaps 4;

QY 21 KTFNGTEGRNSQPMQVGLFEGTSLRCGGVLDHRLVLTAAHCSGRVYVRLGHSLSQL 80
DQ 8 KIVGVTCAANSIPYQVSLNSGSHF--CGGSLINSQWVSAARCYKSRIQVRLGHNIDVL 66
QY 81 DWTEQIRHSGFVTHPGYLGASTSHEDHLRLRLRPVAVTSVQVPLPNDCAATGTC 140
DQ 67 EGNEQFINAKIITHPNFNG--NTLNDMLKLSPATINSRVATVSLPRCAAGTGC 124
QY 141 HVSNGWITNHPNPPDLLQCLNLSIVSHATGCHGVYVPGRTSNMVCAGGVP--GQDACQGD 199
DQ 125 LLSGNGTKSSGSSYPSSLQCLKAPVLSDSCKSSYPGQITGNMVCVFLGCKDSCQGD 184
QY 200 SGGPLVCGVGLQGLVSWGSGVPGCGQDIPGVYTYCKYVDWIRMNRNN 248
DQ 185 SGGPVVCGQLQGVSWGY--GCAQKNKPGVYTKVCNVYVNIQOTIAAN 231

RESULT 5
A35871
trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 04-Mar-2000
C:Accession: A35871; S12117
R:Shi, Y.B.; Brown, D.D.
Genes Dev. 4, 1107-1115, 1990
A:Title: Developmental and thyroid hormone-dependent regulation of pancreatic genes in X
A:Reference number: A35871; MUID:91007255; PMID:2210372
A:Accession: A35871
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-243 <SHI>
A:Cross-references: EMBL:X53458; NID:965162; PIDN:CAA37538.1; PID:965163
C:Superfamily: trypsin, trypsin homology
C:Keywords: hydrolase; protein digestion; serine proteinase
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-20/Domain: activation peptide #status predicted <APT>
F:21-236/Domain: trypsin homology <TRY>
F:24-243/Product: trypsin I #status predicted <MAT>
F:27-157,45-61,129-230,136-203,168-182/Disulfide bonds: #status predicted
F:60,104,197/Active site: His, Asp, Ser #status predicted

Query Match 40.4%; Score 555; DB 2; Length 243;
Best Local Similarity 45.7%; Pred. No. 4.3e-38;
Matches 113; Conservative 42; Mismatches 82; Indels 10; Gaps 6;

QY 6 FLLCLVLGSLQAAT---PKIFNGTEGRNSQPMQVGLFEGTSLRCGGVLDHRLVLTAAH 62
DQ 3 FLLCLVL-LGAAAFDDDKIIGCATCAKSSVPVIVSLNSGYHF--CGGSLITNQWVSAAH 60
QY 63 CSSRVTWRLGHSLSQLDWTEQIRHSGFVTHPGYLGASTSHEDHLRLRLPVRVTS 122
DQ 61 CYKASIQVRLGENIALSEGTGEQFISSSKVIHSGY--NSYITLNDMLKLSPASLNA 118
QY 123 SVQPLPNDCAATGCHVSGWGTINHPNPPDLLQCLNLSIVSHATGCHGVYVPGRTS 182
DQ 119 AVNTVFLPSCSAAGTSCILSHGNTLSNGSNTPDLLQCLNAPILTNACNSAYFGEITA 178
QY 183 NMVCAGGVP--GQDACQGDSSGGPLVCGVGLQGLVSWGSGVPGCGQDIPGVYTYCKYVDWI 241
DQ 179 NMICVGMGEGKDSQCGDSGGPVVVCNQLQGVWSGY--GCAMRNYPGVYTKVCNVNAMI 236
QY 242 RMIMNRNN 248
DQ 237 QNTIAAN 243

RESULT 6
```

```
A53968
serine proteinase SCCE precursor - human
N:Alternate names: stratum corneum chymotryptic enzyme
C:Species: Homo sapiens (man)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999
C:Accession: A53968
R:Hansson, L.; Stromqvist, M.; Baackman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud
J. Biol. Chem. 269, 19420-19426, 1994
A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic e
A:Reference number: A53968; MUID:94308225; PMID:8034709
A:Accession: A53968
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-253 <HAN>
A:Cross-references: GB:L33404; NID:9521214; PIDN:AAC37551.1; PID:9532504
C:Genetics:
A:Gene: GDB:PRSS6; SCCE
A:Cross-references: GDB:377730
A:Map position: 7q35-7q35
C:Superfamily: trypsin; trypsin homology
F:30-245/Domain: trypsin homology <TRY>

Query Match 40.4%; Score 555; DB 2; Length 253;
Best Local Similarity 44.6%; Pred. No. 4.5e-38;
Matches 115; Conservative 37; Mismatches 90; Indels 16; Gaps 6;

QY 1 MGLSIFLLLCVGLS-----QAATPKIFNGTEGRNSQPMQVGLFEGTSLRCGGVLI 52
DQ 1 MARSULLPLQILLLSLALETAGEEAQGDKIIDGAPCARGSHPWQVALLSGNQLHCGGVLV 60
QY 53 DHRWVLTAAHCSGRVYVRLGHSLSQLDWTEQIRHSGFVTHPGYLGASTSHEDHLRL 112
DQ 61 NERWVLTAAHCKXNXYVHLGSDTLG--DRRAQRIKASKSFRHPGY--STQTHVNDMLV 116
QY 113 RLRLPVRVTSVQPLPNDCAATGCHVSGWGTINHPNPPDLLQCLNLSIVSHATC 172
DQ 117 KLSNQARLSSMWKVLPSRCPPGTTCTVSGWGTTSPEVTFPSDLMCVDVKLLSPQDC 176
QY 173 HGYPGRITSNMVCAGGVP--QDACQGDSSGGPLVCGVGLQGLVSWGSGVPGCGQDIPGV 230
DQ 177 TKVYKDLLENMLCA--GIPDSKKNACNGSGGLVCRGTLOGLVSWGTF--PCGQPNPDGV 234
QY 231 YTVICKYVDWIRMNRNN 248
DQ 235 YIQVCKFTKWIINDTKKH 252

RESULT 7
B25528
trypsin (EC 3.4.21.4) precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C:Accession: B25528
R:Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
Nucleic Acids Res. 14, 8307-8330, 1986
A:Title: Sequence organisation and transcriptional regulation of the mouse elastase
A:Reference number: A93646; MUID:87066713; PMID:3641189
A:Accession: B25528
A:Molecule type: mRNA
A:Residues: 1-246 <STE>
A:Cross-references: GB:X04574; NID:954918; PIDN:CAA28243.1; PID:954919
C:Superfamily: trypsin; trypsin homology
C:Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-239/Product: trypsin #status predicted <MAT>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 40.2%; Score 553; DB 2; Length 246;
Best Local Similarity 44.7%; Pred. No. 6.3e-38;
Matches 113; Conservative 45; Mismatches 81; Indels 14; Gaps 7;
```


A; Cross-references: GB:M11599; NID:G164094; PIDN:AAA30899.1; PID:G164095
 C; Superfamily: trypsin; trypsin homology
 C; Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
 F; 1-15/Domain: signal sequence #status predicted <SIG>
 F; 16-23/Domain: activation peptide #status predicted <APT>
 F; 24-247/Product: trypsin, anionic #status predicted <ENZ>
 F; 24-239/Domain: trypsin, homology <TRY>
 F; 30-160, 48-64, 132-233, 339-206, 171-183/Disulfide bonds: #status predicted
 F; 63, 107, 200/Active site: His, Asp, Ser #status predicted
 F; 75, 77, 80, 85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 39.7%; Score 546; DB 1; Length 247;
Best Local Similarity 45.6%; Pred. No. 2.4e-37;
Matches 113; Conservative 38; Mismatches 85; Indels 12; Gaps 6;

Qy	7	LLLCVLGLSQAATP-----KIFNGTECGRNSOPWQVGLFECTSLRCGGVLIDHRSWLTA	61
Dd	5	LILAFLG-AAVATPDDDDKLVGGTYCEENSVP-QVSLNAGYHF-CGSLISDQVVWSAA	62
Qy	62	HCSGRVYVRLGEHSLSOLDWTEQIRHSGFSFVTHPGYLGASTSHEDLRLLRLLPVRVT	121
Dd	63	HCYKSRIQVRLGEYNIDVLEGNQFINSAKVIRHPNY--NSWILDNDIMLIKLSPPAVLN	120
Qy	122	SSVOPLPLPNDCATAGTETCHVSXGWGITMHRNPFPDLLQCINLSIVSHATCHGVYPGEIT	181
Dd	121	ARVATISLPRACAPGTQCLISGMGNLTSSGNTNPPELLQCLUDAPILTOAQCEASYPGOIT	180
Qy	182	SNMWCAAGVGP-GQDACQGDSDGGLPCGGVLOQLVGSWGSPCGQDGIGFYGYTYICKYVDW	240
Dd	181	ENMICAGLEGKGKSCQDGGSPVCNSELQGIYSWGY--GCAQKNKGSVYTKVCNFVDW	238
Qy	241	IRMIMRN	248
Dd	239	IQTIAAN	246

RESULT 11

```

TRDGC
trypsin (EC 3.4.21.4) precursor, cationic - dog
N:Alternate names: cationic trypsinogen
C:Species: Canis lupus familiaris (dog)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: B26273
R:Pinsky, S.D.; LaForge, K.S.; Scheele, G.
Mol. Cell. Biol. 5, 2669-2676, 1985
A:Title: Differential regulation of trypsinogen mRNA translation: full-length
A:Reference number: A26273; MUID:86284628; PMID:3841794
A:Accession: B26273
A:Molecule type: mRNA
A:Residues: 1-246 <PIN>
A:Cross-references: GB:M11590; NID:gl64096; PID:AAA30900.1; PID:gl64097
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-23/Domain: activation peptide #status predicted <APT>
F:24-246/Product: trypsin, cationic #status predicted <ENZ>
F:24-239/Domain: trypsin homology <TRI>
F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

```

Query Match 39.7%; Score 545; DB 1; Length 246;
Best Local Similarity 44.2%; Pred. No. 2.9e-37;
Matches 111; Conservative 43; Mismatches 87; Indels 10; Gaps 6

QY	3	LS	F	L	L	L	C	V	L	S	O	A	---	A	T	P	K	I	F	N	G	T	E	C	R	N	S	Q	P	W	O	V	G	L	P	E	G	T	S	L	R	C	G	V	L	D	H	R	W	T	L	58				
Db	1	M	K	T	F	I	F	A	L	L	G	A	T	A	P	I	D	D	D	K	I	V	G	Y	T	C	S	R	N	S	V	Y	Q	V	S	L	N	S	G	Y	H	F	-	C	G	S	L	I	N	S	Q	W	W	59		
QY	59	T	A	A	H	C	S	G	R	Y	V	W	R	L	G	H	S	L	S	Q	L	D	T	Q	I	R	H	S	G	F	S	V	T	H	P	G	Y	L	G	A	S	T	S	H	E	H	D	L	R	L	R	L	P	V	118	
Db	60	S	A	A	H	C	Y	K	S	R	I	O	V	R	L	G	H	S	L	S	Q	L	D	T	Q	I	R	H	S	G	F	S	V	T	H	P	G	Y	L	G	A	S	T	S	H	E	H	D	L	R	L	R	L	P	V	117

	Y	Q _y	119	RVTSSVQPPLPNDKATAGTCHEVSWGVIINHRNPPDLLQLCLNLISIVSHATCHGVYPG	178
	D _b		118	TLSNRSAIAALPKSCPAAGTQCILSGWNTQTSGONYPDVLQCLKAPILSDSVCRNAYPG	177
	Y	Q _y	179	RITSNMVCAGVP- GDACCGDGGSPLVCGGVLQGLVSWGSVGPCGDGIPIGVITYICKY	237
	D _b		178	QIESNNMCLGYRGGKDSCDGSGPVVNCNGELQGVVWSGA--GCAQKGFGVSPKVCKY	235
	Y		238	VDTIRMIMRN 248	
	D _b		236	VSWICOTIAAN 246	

RESULT 12

A27547
trypsin (EC 3.4.21.4) precursor, cationic - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C:Accession: A27547
R:Pletcher, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C.
Biochemistry 26, 3081-3086, 1987
A:Title: Isolation and characterization of a cDNA encoding rat cationic trypsin
A:Reference number: A27547; MUID:87271f609; PMID:3607011
A:Accession: A27547
A:Molecule type: mRNA
A:Residues: 1-247 <FLB>
A:Cross-references: GB:M16624; NID:G206498; PIDN:AAA41985.1; PID:G206499
C:Superfamily: trypsin; trypsin homology
C:Keywords: calcium binding; hydrolyase; protein digestion; serine proteinase
C:Keywords: calcium binding; hydrolyase; protein digestion; serine proteinase
F:325-240/Domain: trypsin homology <TRY>
F:31-161,49-65,133-234,140-207,172-186/Disulfide bonds: #status predicted
F:64,108,201/Active site: His, Asp, Ser #status predicted
F:76,78,81,86/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 39.6%; Score 544.5; DB 2; Length 247;

Best Local Similarity	43.3%;	Pred No. 3.1e-37;	
Matches 109:	Conservative	43:	Mismatches 89:
Indels	11:	Gaps	5:

3LSIFLLCVLGLSOAA-----TPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWV 57

A vertical ruler scale showing measurements from 0 to 9 inches. The markings are as follows:

Inch	Centimeter
0	0
1	2.54
2	5.08
3	7.62
4	10.16
5	12.70
6	15.24
7	17.78
8	20.32
9	22.86

Db 1 MKALIFLAFLGAVALPLDDDDDKIVGGYTCQKNSLPYQVSLNAGYHF-CGGSLNSQWV 59

Qr 58 I T A A H C S C S R V W V B I G E H S I S O L D W T E O T P H G S E S Y T H P G Y I G A S T S H E H D I R I I R I R P 117

QY 36 11AFHCS93K1WAKDGEH3CQJWMTQ1K1NGDVA THE C10F0A T0M1H1D1K1N1E1K1

DB 60 VSAHCYKSRIQVRLGEHNIDVVEGGEQFIDAKIIRHPSY--NANTFDNDIMLIKLNPS 117

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qy 118 VRVTSSVQPLPLPNDCATAGTECHVSGWGITNHPNPFDDLQCLNLSIVSHATCHGVYF I 177

118 ATLNRSRVSTVSLPRSCGSSGTKCLVSGWGNTLSGGTNPYPSLLOCLDAPVLSDSCKSSYP 177

CTT
CTT

QY 178 GRITSNMVCAGGVP-GQDACQDGGPLVCGGVQLQGLVSWGSLVGPCGQDGIPGVVITYICK 236

[illegible]

DB I/78 GRI.TSNMFCTGFLEGGKDSQCQGDSSGFPVVCNGDQGVVSWGIF--GCAQKNGKFGVITAKVCN 233

Qv 237 YVDWIRMIRNN 248

 • •
 • •

 • •

Db 236 YVNWIQQTVAA 247

RESULT 13

TRRT1

trypsin (EC 3.4.21.4) I precursor - rat

N;Alternate names: trypsinogen I

C;Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence revision 17-Dec-1982 #text change 24-Sep-1999

C:Date: 17-Dec-1982 #sequence_revision: 17 Dec 1982 #seqs_chnge: 17 Dec 1982
C:Accession: B22657; A00948

R; Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.

J. Biol. Chem. 259, 14255-14264, 1984

A;Title: Structure of two related rat pancreatic trypsin genes.
A.Reference number: A22657. MUID:85054880. PMID:6094547

A;Reference Number: A22657; NCID:65034880; PMID:3094347
A:Accession: B22657

A;Molecule type: DNA

A;Residues: 1-246 <CRA>

Search completed: June 16, 2004, 20:34:43
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2004, 20:30:01 ; Search time 45 Seconds
(without alignments)
1738.856 Million cell updates/sec

Title: US-10-015-385A-194
Perfect score: 1374
Sequence: 1 MGSLIFLLCVGLSQAATP.....GVYTYICKYVDWIMRMNN 248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:.*
1: sp_archea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mmc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_virus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	952.5	69.3	234	11 Q9CV76	Q9CV76 mus musculus
2	627.5	45.7	260	4 Q8IW69	Q8IW69 homo sapien
3	608	44.3	254	11 Q8CGR4	Q8CGR4 mus musculus
4	601	43.7	275	4 Q8IXD7	Q8IXD7 homo sapien
5	597.5	43.5	249	11 Q9QYNA	Q9QYNA mus musculus
6	597.5	43.5	276	11 Q9QYN3	Q9QYN3 m hipposas
7	595	43.3	255	4 Q9GRQ0	Q9GRQ0 homo sapien
8	588.5	42.8	293	11 Q9D140	Q9D140 mus musculus
9	588	42.8	276	11 Q8CGR6	Q8CGR6 mus musculus
10	586	42.6	250	11 Q8CGR5	Q8CGR5 mus musculus
11	578	42.1	242	11 Q8OVSA	Q8OVSA mus musculus
12	563	41.0	248	13 Q7SZT1	Q7SZT1 xenopus lae
13	555	40.4	243	13 Q7SZ06	Q7SZ06 xenopus lae
14	553.5	40.3	246	11 Q88301	Q88301 mus musculus
15	553.5	40.3	253	11 Q91Y82	Q91Y82 mus musculus
16	550.5	40.1	247	11 Q9D7Y7	Q9D7Y7 mus musculus

17	549.5	40.0	247	11 Q9CPN9	Q9CPN9 mus musculus
18	546.5	39.8	251	11 Q54854	Q54854 rattus norv
19	544	39.6	253	4 Q8NSN9	Q8NSN9 homo sapien
20	541.5	39.4	247	11 Q9CPN7	Q9CPN7 mus musculus
21	538.5	39.2	249	11 Q91VE3	Q91VE3 mus musculus
22	536	39.0	246	11 Q9ROT7	Q9ROT7 mus musculus
23	536	39.0	246	11 Q7TT42	Q7TT42 mus musculus
24	533	38.8	246	11 Q9ZLR9	Q9ZLR9 mus musculus
25	532	38.7	246	11 Q9QDK9	Q9QDK9 mus musculus
26	525	38.2	240	13 Q98TH0	Q98TH0 engraulis j
27	522	38.0	244	13 Q8QGW3	Q8QGW3 anguilla ja
28	514.5	37.4	242	13 Q7TIR8	Q7TIR8 pangasius h
29	514	37.4	239	11 Q63275	Q63275 rattus norv
30	513.5	37.4	243	13 Q8AV83	Q8AV83 brachydanio
31	513.5	37.4	261	6 Q9NIQ1	Q9NIQ1 sagunius oe
32	513.5	37.4	278	11 Q99M20	Q99M20 mus musculus
33	511.5	37.2	261	6 Q29474	Q29474 canis famil
34	510.5	37.2	238	13 Q9W7Q6	Q9W7Q6 paralicthy
35	509.5	37.1	222	13 Q8AV11	Q8AV11 oncorhynch
36	509.5	37.1	235	11 Q63274	Q63274 rattus norv
37	508	37.0	242	13 Q9W7Q7	Q9W7Q7 paralicthy
38	507	36.9	244	13 Q42159	Q42159 petromyzon
39	506	36.8	245	13 Q42160	Q42160 petromyzon
40	505	36.8	242	13 Q7SX90	Q7SX90 brachydanio
41	505	36.8	249	13 Q92046	Q92046 disostichu
42	504.5	36.7	269	4 Q8IU55	Q8IU55 homo sapien
43	503	36.6	241	13 Q98TG9	Q98TG9 engraulis j
44	502	36.5	249	13 Q9W6K0	Q9W6K0 notothenia
45	501.5	36.5	237	13 Q91515	Q91515 fugu rubrip

ALIGNMENTS

RESULT 1

Q9CV76	PRELIMINARY;	PRT;	234 AA.
AC Q9CV76			
DT 01-JUN-2001 (TRENBLrel. 17, Created)			
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)			
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)			
DE 2310008B01R1K protein (Fragment).			
GN 2310008B01R1K.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Tongue;			
RX MEDLINE=21085660; PubMed=11217851;			
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,			
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,			
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombærce P.,			
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA Hayashizaki Y.;			
RT "Functional annotation of a full-length mouse cDNA collection.";			
RL Nature 409:685-690(2001).			
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
DR EMBL; AK009217; BAB26143.1; -.			

```
DR HSP; P00763; IDPO.
DR MEROPS; S01.020; -.
DR MGD; MGI:1916761; 2310008B01Rik.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 234 AA; 25888 MW; 6D81E609EDD39110 CRC64;

Query Match 69.3%; Score 952.5; DB 11; Length 234;
Best Local Similarity 70.6%; Pred. No. 4.1e-84;
Matches 166; Conservative 27; Mismatches 41; Indels 1; Gaps 1;

QY 14 LSOAATPKIENGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAHCSGRYVRLG 73
Db 1 LSQLADREKINGVECVKNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAHCR-DKYVYVRLG 59
QY 74 EHSLSQDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLRVTSVQPLPLPND 133
Db 60 EHSITKLDWTQIRHSGFSVTHPGYLGASTSHEHDLRLRLPLRVTSVQPLPLPND 133
QY 134 ATAGTECHVSGWGITNHPNPPFDLLQCLNLSIVSHATCHGVYVPGRI 193
Db 120 VTTGAMCHVSGWGITNHPNPPFDLLQCLNLSIVSHATCHGVYVPGRI 193
QY 194 DACQGSGLVCGVQLGVLSVSGVPCGQDGPVYVYICKYVDVIRIMINN 248
Db 180 DACQGSGLVCGVQLGVLSVSGVPCGQDGPVYVYICKYVDVIRIVIRN 234

RESULT 2
Q81W69 PRELIMINARY; PRT; 260 AA.
AC Q81W69
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Srausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040887; AA040887.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein.
SQ SEQUENCE 260 AA; 28090 MW; EF5934EB96295660 CRC64;

Query Match 45.7%; Score 627.5; DB 4; Length 260;
Best Local Similarity 50.2%; Pred. No. 1.4e-52;
Matches 123; Conservative 24; Mismatches 93; Indels 5; Gaps 3;

QY 5 IFLL--CVLGLSOAATPKIENGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAA 61
Db 13 MFLLLGAGWAGHSGRAQEDKVLGGHECPHSPQWQAALFQGGQLCGVLVGNWVLTAA 72
QY 62 HCSGSRYVRLGHSLSQLDWTQIRHSGFSVTHPGYLGAST-SHEHDLRLRLPLRV 120
Db 73 HCKPKYTVRLGHSLSQLDWTQIRHSGFSVTHPGYLGAST-SHEHDLRLRLPLRV 132
QY 121 TSSVQPLPLPNDCATAGTECHVSGWGITNHPNPPFDLLQCLNLSIVSHATCHGVYVPGRI 180
Db 133 GSKVKPISLADHCTQPGQKCTVSGWGTVTSPPNFFDTLNCABVKIFPKKCBDAVFGOI 192
QY 181 TSNMVCAGVPGQDACQGSGLVCGVQLGVLSVSGVPCGQDGPVYVYICKYVDW 240
Db 193 TDVMVCAGSKGADTCQGSGLVCGVQLGVLSVSGVPCGQDGPVYVYICKYVDW 251
QY 241 IRMIM 245
Db 252 IKXII 256

RESULT 3
Q8CGR4 PRELIMINARY; PRT; 254 AA.
AC Q8CGR4
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prostin.
GN KKL15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Leary A.Y.; Lundwall A.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL152434; AN78422.1; -.
DR MGD; MGI:2447533; Kkl15.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 254 AA; 28042 MW; AA9E38BEDD01861 CRC64;

Query Match 44.3%; Score 608; DB 11; Length 254;
Best Local Similarity 49.0%; Pred. No. 1e-50;
Matches 124; Conservative 30; Mismatches 77; Indels 22; Gaps 7;

QY 5 IFLLCVLGLSOAAT--PKIENGTECGNSQPMQVGLFEGTSLRCGGVLIHRWVLTAAH 62
Db 1 MWLLAFVLLVSAQDGKVLGEGECVPHSQVQALFEGRFGNCGAFLISPRWVLTAAH 60
```

```
QY 63 CSGSRVYVRLGHEHSLDQWTEQIRHSGFSVTHPGYLGASTSHEDHRLRLPLRVTS 122
DB 61 CQTRFMRVRLGHEHSLDQWTEQIRHSGFSVTHPGYLGASTSHEDHRLRLPLRVTS 118
QY 123 SVQPLPLPNDCATAGTECHVSGWGTINHPRNP-----FPDLQLCLNLSIVSHAT 171
DB 119 YVRPVALPRCPGIGEDCVVSGWGLLS-DNNPGATGSKSHVRUPDTLHCANISIISEAS 177
QY 172 CHGVYPRGRIITSNMVCAG---GGVPGQDACQDGGGGLVCGGVLQGLVSGVSGVPGQDGP 228
DB 178 CNKDYPRGRLVPLTMVCAGVEGG---GTDSCGDSGGGLVCGGALQGLVSGVGV-PCDITTKP 234
QY 229 GVVTYICKYVDWI 241
DB 235 GVVTKVCYSLEWI 247

RESULT 4
QY Q8IXD7 PRELIMINARY; PRT; 275 AA.
AC Q8IXD7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE variant form hippostasin/KLK11.
GN KLK11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Nakamura T., Mitsui S., Miki T., Yamaguchi N.;
RT "Molecular cloning and expression of a variant form of
RT hippostasin/KLK11 in prostate.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078780; BAC34105.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 275 AA; 30165 MW; 257A42B28F40E2C4 CRC64;

Query Match 43.7%; Score 601; DB 4; Length 275;
Best Local Similarity 44.2%; Pred. No. 5.5e-50;
Matches 121; Conservative 35; Mismatches 88; Indels 30; Gaps 5;

QY 3 LSIFLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTAH 62
DB 4 LQILALATGLVGET-RIKGFCKPHSQPWQAALFEKRLCGATLIAPRWLLTAH 62
QY 63 CSG-----SRYVRLGHEHSLDQWTEQIRHSGFSVTHPG 97
DB 63 CLKPWVLSLTPHVSPLSSNVCVLSHLSRYIVHLGQHLKEEGCEQRTATFPPHP 122
QY 98 YLGA--STSHEDHRLRLPLRVTSVQPLPNDCATAGTECHVSGWGTINHPRNP 155
DB 123 FNNSLPNKDRNDIMLVQWASPVSIWVAPLTLSSRCVYTAGTSLISGWSGTSPPQLRL 182
QY 156 PDLLQCLNLSIVSHATCHGVYPRGRIITSNMVCAG-GVPGQDACQDGGGGLVCGVQLV 214
DB 183 PHTLRCANITITIEHQKCNAYPGNITDTWVCASVQEGGKDSQDGGGLVCGVQLV 242
```

```
QY 215 SWGSGVPGQDGIPIGVYTYICKYVDWIRMNRN 248
DB 243 SWGQ-DECAITRPGVYTYICKYVDWIQETMKN 275

RESULT 5
QY Q9QYN4 PRELIMINARY; PRT; 249 AA.
AC Q9QYN4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hippostasin.
GN PRSS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "cDNA cloning of a novel brain serine protease, Hippostasin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016226; BAA88825.1; -.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.257; -.
DR MGD; MGI:192977; Prss20.
DR GO; GO:0005576; C:extracellular; IDA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 249 AA; 27604 MW; F9FF9CB457D727D5 CRC64;

Query Match 43.5%; Score 597.5; DB 11; Length 249;
Best Local Similarity 47.0%; Pred. No. 1.1e-49;
Matches 118; Conservative 33; Mismatches 95; Indels 5; Gaps 4;

QY 1 MGLSIFLLCVLGSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWLTA 60
DB 1 MLRLIALALVTGVHGET-RIKGYECRPHSQPWQVAFQKTRLCCGATLIAPKWLTA 59
QY 61 AHCSGSRYVRLGHEHSLDQWTEQIRHSGFSVTHPGYLG--STSHEDHRLRLPLRV 118
DB 60 AHCRPHYVILLGHEHSLDQWTEQIRHSGFSVTHPGYLG--STSHEDHRLRLPLRV 119
QY 119 RVTSSVQPLPLPNDCATAGTECHVSGWGTINHPRNPFFDLQLCLNLSIVSHATCHGVY 178
DB 120 FFTRAVQPLTLSPHCVAAGTSLISGWTTSQRLRPHSLRCANVSIIEHKECEKAYPG 179
QY 179 RTTSMNVCAG-GVPGQDACQDGGGGLVCGGVLQGLVSGVSGVPGQDGIPIGVYTYICKY 237
DB 180 NITDTMLCASVKEGKDSQDGGGLVCGVQLVCGVQLVCGVQLVCGVQLVCGVQLV 238
QY 238 VDWIRMNRN 248
DB 239 FNWIHEVMNRN 249

RESULT 6
QY Q9QYN3 PRELIMINARY; PRT; 276 AA.
ID Q9QYN3
```

```

AC  DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DE  DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE  DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  DE Hippotasin prostate type (Adult male tongue cDNA, RIKEN full-length
DE  DE enriched library, clone:2310015108 product:protease, serine, 20, full
DE  DE insert sequence) (Adult male tongue cDNA, RIKEN full-length enriched
DE  DE library, clone:2310040F07 product:protease, serine, 20, full insert
DE  DE sequence).
DE  DE PRSS20.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RN  SEQUENCE FROM N.A.
RP  TISSUE=Prostate;
RC  TISSUE=Prostate;
RX  MEDLINE=20525460; PubMed=11072089;
RA  Mitsui S., Okui A., Kominami K., Uemura H., Yamaguchi N.:
RT  "cDNA cloning and tissue-specific splicing variants of mouse
RL  hippotasin/TLSP (PRSS20).";
RN  Biochim. Biophys. Acta 1494:206-210(2000).
[3]
RN  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Tongue;
RA  Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA  Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA  Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA  Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA  Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA  Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA  Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA  Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA  Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA  Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA  Muramatsu M., Hayashizaki Y.;
RL  Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Tongue;
RA  The FANTOM Consortium;
RA  the RIKEN Genome Exploration Research Group Phase I & II Team;
RA  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RN  Nature 420:563-573(2002).
[5]
RN  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Tongue;
RA  RIKEN FANTOM Consortium;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RN  Nature 409:685-690(2001).
[6]
RN  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Tongue;
RX  MEDLINE=20499374; PubMed=11042159;
RA  Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA  Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT  "Normalization and subtraction of cap-trapper-selected cDNAs to
RN  prepare full-length cDNA libraries for rapid discovery of new genes.";
RL  Genome Res. 10:1617-1630(2000).
[8]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Tongue;
RX  MEDLINE=20530913; PubMed=11076861;
RA  Shibata K., Itoh M., Aizawa K., Katsunai T., Tashiro H., Itoh M.,
RA  Konno H., Akiyama J., Nishi K., Katsunai T., Nishine T., Harada A.,
RA  Sumi N., Ishii Y., Nakamura S., Hazama M., Ikegami T., Kashiwagi K.,
RA  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA  Fujiwaka S., Inoue K., Togawa Y., Izawa K., Tanaka T., Matsura S., Kawai J.,
RA  Yoneda Y., Ishikawa T., Ogawa K., Tanaka T., Matsura S., Kawai J.,
RA  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT  "RIKEN integrated sequence analysis (RISA) system-384-format
RN  sequencing pipeline with 384 multicapillary sequencer.";
RL  Genome Res. 10:1757-1771(2000)
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR  EMBL; AB016227; BAB36955.1; -.
DR  EMBL; AK009360; BAB26241.2; -.
DR  EMBL; AK009720; BAB26461.2; -.
DR  HSSP; P00763; IDPO.
DR  MEROPS; S01.257; -.
DR  MGD; MGI:1929977; Prss20.
DR  GO; GO:0005576; C:extracellular; IDA.
DR  InterPro; IPR009003; Cys_Ser_trypsin.
DR  InterPro; IPR001254; Peptidase_S1.
DR  InterPro; IPR001314; Peptidase_S1A.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  SMART; SM00020; Tryp_Src; 1.
DR  PROSITE; PS0240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
KW  Hydrolase; Protease; Serine protease.
SQ  SEQUENCE 276 AA; 30753 MW; 90BDC03A8AB178D6 CRC64;

Query Match 43.5%; Score 597.5; DB 11; Length 276;
Best Local Similarity 47.0%; Pred. No. 1.2e-49;
Matches 118; Conservative 33; Mismatches 95; Indels 5; Gaps 4;

Qy 1 MGLSIFLLLVGLSQATPKIFNGTECGRNSOPWQVGLFEGTSLRCGVLIDHRWLTA 60
Db 28 MLRLIALALVTGVTGET-R-I-KGYECRPHSQPWQVAFKTRLLCGATLIAPKWLTA 86
Qy 61 AHCSGSRYWVRLGHSLSQLDWTEQIRHSGFVSTHPGYLGA--STSHEDLRLRLPLV 118
Db 87 AHCRKPHVILLGEHNLEKTDGCEQRMMATESPFPDFNNSLPNKDHNDMLVKMSPV 146
Qy 119 RVTSVQPLPLPNDACATAGTECHVSGWGIHNPENPFDDLLQCLNLSIVSHATCGVYPG 178
Db 147 FFTRAVQPLTLSPFCVAAAGTSLISGWGTTSSPOLRPLHSLRCANVSIIEHKECEKAYPG 206
Qy 179 RITSNMVCAG-GVPGDQACQDGGPLVCGVGLQGLVSWGSGVPGCGQDGPVGYTYCKY 237
Db 207 NIITDMLCASVRKESGKSCQDGGPLVCGVGLQGLIISWGQ-DPCAIVTRKPGVTVKCKY 265
Qy 238 VDWTRMIMRNN 248
Db 266 FNWIEHVMRNN 276

RESULT 7
Q96RQ0 ID Q96RQ0 PRELIMINARY; PRT; 255 AA.
AC Q96RQ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prostinogen.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21226193; PubMed=11327827;

```

```
RA Takayama T.X., Carter C.A., Deng T.;
RT "Activation of prostate-specific antigen precursor (pro-PSA) by
RT prostin, a novel human prostatic serine protease identified by
RT degenerate PCR";
RL Biochemistry 40:1679-1687(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF303046; AAK62813.1; -.
DR HSSP; P00761; IAN1.
DR MEROPS; S01.081; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 255 AA; 27986 MW; 00D5B79E14B9468F CRC64;

Query Match 43.3%; Score 595; DB 4; Length 255;
Best Local Similarity 46.5%; Pred. No. 1.9e-49;
Matches 119; Conservative 32; Mismatches 87; Indels 18; Gaps 6;

QY 8 LLVLGLGQAAT----PKIFNTEGRNSQPMQVGLFECTSLRCGVLIDHRWVLTAAHC 63
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 LLLTSLFLASTAQDQKLEDECAHPSQPMQVGLYERGNCGASLSPHWVLSAAHC 62
QY 64 SGRYVWRLGHSLSQLDWTEQIRHSGFVTHPGYLGASTSHEHDLRLRLPVRVTSS 123
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 QSRFVRLGHNLRKRDGPEQLRTTSRVIPHRY--EASRNDIMLLRLVQPARLNQ 120
QY 124 VQPLPLNDCAATAGTECHVSGMGTNH-----PRN--PPDLLOCLNLSIVSHATCH 173
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 VRPAVLPTCRPHGACVVGSLVSHNEPFGTAGSPRSQVSLPDLTLCANIGIISDTS 180
QY 174 GYVPGRTSNMVCAGV-PGODACDGSGLVCGVLQGLVSGVSGPCGDGIPGVYT 232
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 KYPRLNTMVCAGAEGRGAECSGDSGGLVCGILQGIIVSGDV-PCDNTTREGVYT 239
QY 233 YICKYVDWIRMIMRN 248
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 KVCHYLEWIRETKRN 255

RESULT 8
Q9D140
ID Q9D140 PRELIMINARY; PRT; 293 AA.
AC Q9D140;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE L110030019Rik protein.
GN L110030019Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
```



```
Query Match 42.1%; Score 578; DB 11; Length 242;
Best Local Similarity 49.1%; Pred. No. 7.9e-48;
Matches 114; Conservative 35; Mismatches 75; Indels 8; Gaps 5;

Qy 21 KIENGTEGRNSQPMQVCLFGTSLR--CGVLLDHRWLVTAHCGSGRYVRLGHSLS 78
Db 15 KIIIGYCVRSQPMQVCLFGTSLR--CGVLLDHRWLVTAHCGSGRYVRLGHSLS 74
Qy 79 QLDWTEQIRHSGFSVTHPGYLCASTSHEHDLRLRLRLPVRVTSVQPLPNDCAATAGT 138
Db 75 RWEATQVVRVARQVPHQY--QPOAHNDLMLLKLQKVLGRAVKTVISVASSCASPQT 132
Qy 139 ECHVSGWGITNHRNPPFDLLQCLNLSIVSHATCHGVIPGRITSNMVCAGGVP--GDAC 196
Db 133 PCRVSGWGTIASPIARYPTALQCVNVMIMSEQACHRAYPEIGITSGMVCA-GVPEGGKDCS 191
Qy 197 QGDSGGPLVCGGLVQLGWSVGPVCGQDGPVGVYTKYVDWIRLMIRNN 248
Db 192 QGDSGGPLVCGGLVQLGWSVGPVCGQDGPVGVYTKYVDWIRLMIRNN 242

RESULT 12
Q7SZT1 PRELIMINARY; PRT; 248 AA.
AC Q7SZT1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Trypsin (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner B.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX Klein S., Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RM EMBL; BC056068; AAH56068.1; -.
```

```
FT NON TER 1
SQ SEQUENCE 248 AA; 26494 MW; 1ED9D0D218702860 CRC64;

Query Match 41.0%; Score 563; DB 13; Length 248;
Best Local Similarity 45.0%; Pred. No. 2.3e-46;
Matches 112; Conservative 46; Mismatches 83; Indels 8; Gaps 5;

Qy 3 LSIFLLCIVLGLSQA--ATPKIFNGTEGRNSQPMQVCLFGTSLRCCGVLLDHRWLVTA 60
Db 5 MKELVILVLLGAAVAFEDDDKIVGGFTCAKNAVPPQVSLNAGYHF-CGSLINSQWVVA 63
Qy 61 AHCSGSRVWVLGHSLSQLDWTEQIRHSGFSVTHPGYLCASTSHEHDLRLRLRLPVRV 120
Db 64 AHCYKRSITQVRLGHEHNLNEGTEQFIDSKVIPHNY--NSRLNDNDIMLIKLTSTLRL 121
Qy 121 TSSVQPLPLPNDCAATAGTECHVSGWGITNHRNPPFDLLQCLNLSIVSHATCHGVYPCRI 180
Db 122 SANTQSVLPASACASAGTNCILSGWNTLSSGTWYPDLLQCLNAPILTDSCSNYPGEI 181
Qy 181 TSNMVCAGGVP--GDACQDGGPLVCGGLVQLGWSVGPVCGQDGPVGVYTKYVD 239
Db 182 TKNMFCAGFLAGKDKSCQDGGPVVCGVQGLQGVVSGY--GCAQRNYPGVYTKVCFEVT 239
Qy 240 WIRLMIRNN 248
Db 240 WIQSTISSN 248

RESULT 13
Q7SZ06 PRELIMINARY; PRT; 243 AA.
AC Q7SZ06;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX Klein S., Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RM EMBL; BC056068; AAH56068.1; -.
```



```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RC SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Strausberg R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054194; AAH54194.1; -.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25476 MW; C5B82B458B209F31 CRC64;

Query Match 40.4%; Score 555; DB 13; Length 243;
Best Local Similarity 45.7%; Pred. No. 1.3e-45;
Matches 113; Conservative 42; Mismatches 82; Indels 10; Gaps 6;

Qy 6 FLLLCVLGLSQAAT---PKIFNGTECGRNQFVQVGLFEGTSLRCGGVLIDHRLVLTAAH 62
Db 3 FLLLCVL-LGAAAFDDDKIIGATCAKSSVPYIVSLNSGVHF-CGGSLLTNQWVSAAH 60
Qy 63 CSGSRVYVRLGEHSLDQWTEQIRHSGFVTHPGYLGASTSHHDLRLRLPVRVTS 122
Db 61 CYKASIQVRLGEHSLSEGTEQFISSKVIRHSGY--NSYTLNDIMLIKLSPPASLNA 118
Qy 123 SVQPLPLPNDCATAGTECHVSGWGITNHPNPFDPDLLOCLNLSIVSHATCHGVYVPGRIITS 182
Db 119 AVNTVPLPFGCSAAGTSCLSGWNLTLSNGSNYPDLLOCLNAPILTNACQNSAYPGEITA 178
Qy 183 NMVCAAGVY-PQDACQSDSGPLVCGGVLOGLVSGSVGPGCGQDGIPOGVYTYICKYVDWI 241
Db 179 NMICVGFMEGKDSQCGSDGPGVVCNGQLQGVWSGWY--GCAMRNYPGVYTKVCYNNAWI 236
Qy 242 RMIMRN 248
Db 237 QNTIAAN 243

RESULT 14
O88301 ID O88301 PRELIMINARY; PRT; 246 AA.
AC O88301;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine protease (BSP).
GN PRS18 OR MBSP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Matsui H., Takahashi T.;
RT "Mouse serine protease preferentially expressed in brain.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Skin;
RA Meier N., Dear T.N., Boehm T.;
RT "A novel serine protease expressed in the hair follicle.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RA Matsui H., Takahashi T.;
RT "The sequences of mouse serine protease gene expressed in brain.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB015206; BAA28895.1; -.
DR EMBL; Y18723; CAA7269.1; -.
DR EMBL; AB032402; BAA84544.1; -.
DR HSP; P00763; 1DPO.
DR MEROPS; S01.236; -.
DR MGD; MGI:1343166; Prss18.
DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 246 AA; 27496 MW; E20C080087139B63 CRC64;

Query Match 40.3%; Score 553.5; DB 11; Length 246;
Best Local Similarity 43.5%; Pred. No. 1.9e-45;
Matches 107; Conservative 46; Mismatches 82; Indels 11; Gaps 6;

Qy 7 LLLCVLGLSQA---TPKIFNGTECGRNQFVQVGLFEGTSLRCGGVLIDHRLVLTAAH 62
Db 4 LALLCLV-LAKSAWSEEQKVHGGPCUKDSHPFQAALYTSCHLLCGGVLDIPQWVLTAAH 62
Qy 63 CSGSRVYVRLGEHSLDQWTEQIRHSGFVTHPGYLGASTSHHDLRLRLPVRVTS 122
Db 63 CKENLVQLIGKMLRQTETFRQISVDRITVHPRY--NPETHNDIMVHLKNPVKFSK 120
Qy 123 SVQPLPLPNDCATAGTECHVSGWGITNHPNPFDPDLLOCLNLSIVSHATCHGVYVPGRIITS 182
Db 121 KIQLPLKNDCSBENPNCQLLWG--KXMGDFPDTIQCADVHLVPREQCERAYPGKITQ 178
Qy 183 NMVCAAGVY-PQDACQSDSGPLVCGGVLOGLVSGSVGPGCGQDGIPOGVYTYICKYVDWI 241
Db 179 SMVCAADMKEGNSQCGSDGPGVVCNGQLQGVWSGDM-PCGSKEKPGVYTDVCTHRWI 237
Qy 242 RMIMRN 247
Db 238 QNILRN 243

RESULT 15
Q91Y82 ID Q91Y82 PRELIMINARY; PRT; 253 AA.
AC Q91Y82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurosin (Protease, serine, 18).
GN PRS18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "cDNA cloning and characterization of mouse brain specific serine
protease, Neurosin.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB008928; BAB55605.1; -.
DR EMBL; BC031119; AAH31119.1; -.
DR HSP; P00761; 1ANI.
DR MEROPS; S01.236; -.
DR MGD; MGI:1343166; Prss18.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
```

DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00869; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.

SQ SEQUENCE 253 AA; 28329 MW; C5EF98C7EEF2FBC1 CRC64;

Query Match 40.3%; Score 553.5; DB 11; Length 253;
Best Local Similarity 43.5%; Pred. No. 2e-45;
Matches 107; Conservative 46; Mismatches 82; Indels 11; Gaps 6;

QY	7	LLLCVLGLSQAA----	TPKIFNGTEGRNSQPMQVGLFEGETSLRCGGVLIDHEWVLTAAH	62
DB	11	LALCLV-LAKSAWSEQEKKVHGPGCLKDSHPFQAALYTSGHLLCGGVLDIPQWVLTAAH		69
QY	63	CGSGRYWVRLGEHSLQDLWTEQIRHSGFSVTHPGYLGASTSHSHDLRLRLPLVVRVTS		122
DB	70	CKKPNLQVILGKHLNQTFQFQISVDRIVHPRY--NPETHDNDIMVHLKNPVKESK		127
QY	123	SVQPLPLPNDCAATAGTECHVSGWGIYNHPRNPFPDILLQCLNLSIVSHATCHGVYGRITS		182
DB	128	KIQPLPLKNDCEENPNCQILGWG--KMENGDFPDTIQCADVHLVPREOCERAYPGKITQ		185
QY	183	NMYCAGGV-PGQACOGSGGGLVCGVLOGLVSGSVGPGQGDGIPGVVTVICKYVDWI		241
DB	186	SMVCGDMKEGNDSCQDSGGPLVCGGRRLRGLVSGDM-PCGSKEKPGVTVTDVCTHIRWI		244
QY	242	RMIMRN		247
DB	245	QNILRN		250

Search completed: June 16, 2004, 20:34:11
Job time : 47 secs